

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 10:36:43 ; Search time 73.38 Seconds
(without alignments)
12.109 Million cell updates/sec

Title: US-09-543-188a-1

Perfect score: 55

Sequence: 1 GWGQPHGG 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11107396 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	14	22	Octa-peptide motif
2	55	100.0	18	14	Bovine prion prote
3	55	100.0	18	14	Ovine prion protei
4	55	100.0	18	14	Human prion protei
5	55	100.0	26	14	Prion protein regi
6	55	100.0	42	20	Bovine prion prote
7	55	100.0	178	19	Peptide sequences
8	55	100.0	208	21	Mouse prion protei
9	55	100.0	208	21	Human prion protei
10	55	100.0	208	21	Mouse prion protei
11	55	100.0	208	21	Human prion protei

12	55	100.0	208	22	AA82110	Hamster Prp peptid
13	55	100.0	211	22	AA82110	Amino acid sequenc
14	55	100.0	212	22	AA82110	Cattle prion prote
15	55	100.0	217	21	AA82110	Cattle prion prote
16	55	100.0	217	21	AA82110	Cattle prion prote
17	55	100.0	219	19	AA82110	Bovine prion prote
18	55	100.0	219	20	AA82110	Bovine prion prote
19	55	100.0	245	22	AA82110	Monkey prion prote
20	55	100.0	245	22	AA82110	Monkey prion prote
21	55	100.0	245	22	AA82110	Cercopithe prion p
22	55	100.0	250	22	AA82110	Rabbit prion prote
23	55	100.0	253	17	AA82110	Human prion protei
24	55	100.0	253	19	AA82110	Human prion protei
25	55	100.0	253	20	AA82110	Human prion protei
26	55	100.0	253	20	AA82110	Human prion protei
27	55	100.0	253	21	AA82110	Human prion protei
28	55	100.0	253	21	AA82110	Human prion protei
29	55	100.0	253	22	AA82110	Human prion protei
30	55	100.0	253	22	AA82110	Human prion protei
31	55	100.0	253	22	AA82110	Human prion protei
32	55	100.0	253	22	AA82110	Chimpanzee prion p
33	55	100.0	253	22	AA82110	Orangutan prion p
34	55	100.0	253	22	AA82110	Gorilla prion prote
35	55	100.0	253	22	AA82110	Rhesus monkey prion
36	55	100.0	253	22	AA82110	Gibbon prion prote
37	55	100.0	253	22	AA82110	Prion protein cell
38	55	100.0	253	22	AA82110	Prion protein cell
39	55	100.0	253	22	AA82110	Prion protein cell
40	55	100.0	253	22	AA82110	Prion protein cell
41	55	100.0	253	22	AA82110	Marmoset prion pro
42	55	100.0	253	22	AA82110	Hamadryas prion pr
43	55	100.0	253	22	AA82110	Guereza prion prot
44	55	100.0	253	22	AA82110	Capuchin prion pro
45	55	100.0	253	22	AA82110	Prion protein cell

ALIGNMENTS

RESULT 1

ID AAB84521 standard; peptide; 14 AA.

XX AAB84521;

DT 05-SEP-2001 (first entry)

XX Octa-peptide motif repeated in human prion protein.

DE Prion protein; subacute transmissible spongiform encephalopathy; ESST;
KW Creutzfeld-Jakob disease; bovine spongiform encephalopathy; scrapie.

XX Homo sapiens.

XX WO200135104-A1.

PD 17-MAY-2001.

XX 13-NOV-2000; 2000WO-FR03159.

XX 12-NOV-1999; 99FR-0014242.

XX (COMS) COMMISSARIAT ENERGIE ATOMIQUE.

XX Deslys J, Comoy E, Grassi J;

XX WPI; 2001-408079/43.

PT Diagnosis of prion diseases, by treatment with proteinase K and
PT detecting retention of octapeptide repeat motifs, including
PT differentiation between prion strains -

PS Example 1; Page 13; 51pp; French.

XX The present sequence represents an octa-peptide repeat motif of a
 CC prion protein. The specification describes a method for the diagnosis
 CC of subacute transmissible spongiform encephalopathy (ESST) caused by a
 CC strain of unconventional transmissible agent. The method comprises
 CC detecting abnormal prion protein in a biological sample. The sample is
 CC treated with at least one proteinase K while retaining at least some of
 CC the octapeptide repeats in the prion protein, then treatment with a
 CC ligand for octapeptide repeats and detecting any formation of a complex.
 CC The method is used to diagnose ESST, particularly (new variant)
 CC Creutzfeldt-Jakob diseases, bovine spongiform encephalopathy and
 CC scrapie.

XX Sequence 14 AA;

Query Match 100.0%; Score 55; DB 22; Length 14;
 Best Local Similarity 100.0%; Pred. No. 0.028;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GWGQPHGG 8

Db 2 gwggphgg 9

RESULT 2

AAR38026
 ID AAR38026 standard; protein; 18 AA.

XX AAR38026;

DT 14-OCT-1993 (first entry)

DE Bovine prion protein region E #2.

XX Antigen; prion; protein; region; frame shift; repeat; mutation; PrPc;
 KW FSA; FSB; subfragment; antibody; treatment; spongiform encephalopathy;
 KW human; sheep; cattle; cellular binding; aggregation; mammal; scrapie;
 KW immune system; PrPsc; ratio-inverso peptide; enzymatic degradation;
 KW resistance.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "One or more residues or may be absent"

FT Misc-difference 2 /note= "May be absent"

FT Misc-difference 3 /note= "May be absent"

FT Misc-difference 4 /note= "May be absent"

FT Misc-difference 5 /note= "May be absent"

FT Misc-difference 14 /note= "May be absent"

FT Misc-difference 15 /note= "May be absent"

FT Misc-difference 16 /note= "May be absent"

FT Misc-difference 17 /note= "May be absent"

FT Misc-difference 18 /note= "One or more residue or may be absent"

XX WO9311155-A.

XX 10-JUN-1993.

XX 03-DEC-1992; 92WO-GB02246.

XX 03-DEC-1991; 91GB-0025747.

XX 10-JUL-1992; 92GB-0014663.

XX (PROT-) PROTEUS MOLECULAR DESIGN LTD.

XX Fishleigh RV, Mee RP, Robson B;

XX WPI; 1993-196994/24.

XX New polypeptide(s) contg. antigenic site of prion protein -

XX useful for treatment and diagnosis of mammalian encephalopathies

XX e.g. Creutzfeld-Jacob disease and kuru

XX Claim 23; Page 70; 82pp; English.

XX The sequences given in AAR38025-36 represent polypeptides derived from
 CC an antigenic site, region E, of a prion protein. Prion proteins
 CC comprise six regions of interest (A-F), and two related frame shift
 CC peptides sequences caused by a repeating section in region E having
 CC a nucleic acid coding sequence frame shift mutation of +1 (Fsa) or -1
 CC (FSb) (see also AAR38037-38). These peptides and antibodies raised
 CC against these may be used to treat or prevent spongiform encephalopathy
 CC in humans, sheep or cattle. They can be used to block cellular binding
 CC and aggregation of prion proteins and to stimulate the mammalian immune
 CC system. These peptides may be used to distinguish between the normal
 CC form of prion protein (PrPc) and the scrapie-associated form (PrPsc).
 CC These peptides may include rare or synthetic amino acids or a ratio-
 CC inverso peptide modification to improve resistance to enzymatic
 CC degradation.

XX Sequence 18 AA;

Query Match 100.0%; Score 55; DB 14; Length 18;

Best Local Similarity 100.0%; Pred. No. 0.035;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GWGQPHGG 8

Db 4 gwggphgg 11

RESULT 3

AAR38029
 ID AAR38029 standard; protein; 18 AA.

XX AAR38029;

DT 14-OCT-1993 (first entry)

XX Ovine prion protein region E #2.

XX Antigen; prion; protein; region; frame shift; repeat; mutation; PrPc;
 KW FSA; FSB; subfragment; antibody; treatment; spongiform encephalopathy;
 KW human; sheep; cattle; cellular binding; aggregation; mammal; scrapie;
 KW immune system; PrPsc; ratio-inverso peptide; enzymatic degradation;
 KW resistance.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "One or more residues or may be absent"

FT Misc-difference 2 /note= "May be absent"

FT Misc-difference 3 /note= "May be absent"

FT Misc-difference 4 /note= "May be absent"

FT Misc-difference 5 /note= "May be absent"

FT Misc-difference 14 /note= "May be absent"

FT Misc-difference 15 /note= "May be absent"

```

FT Misc-difference 16 /note= "May be absent"
FT Misc-difference 17 /note= "May be absent"
FT Misc-difference 18 /note= "One or more residue or may be absent."
XX
PN WO9311155-A.
PD 10-JUN-1993.
XX
XX 03-DEC-1992; 92WO-GB02246.
XX
PR 03-DEC-1991; 91GB-0025747.
PR 10-JUL-1992; 92GB-0014663.
XX
PA (PROT-) PROTEUS MOLECULAR DESIGN LTD.
XX
XX Fishleigh RV, Mee RP, Robson B;
PI WPI; 1993-196994/24.
DR
XX
XX New polypeptide(s) contg. antigenic site of prion protein -
PT useful for treatment and diagnosis of mammalian encephalopathies
PT e.g. Creutzfeld-Jacob disease and kuru
XX
XX Claim 23; Page 70-71; 82pp; English.
XX
XX The sequences given in AAR38025-36 represent polypeptides derived from
CC an antigenic site, region E, of a prion protein. Prion proteins
CC comprise six regions of interest (A-F), and two related frame shift
CC peptides sequences caused by a repeating section in region E having
CC a nucleic acid coding sequence frame shift mutation of +1 (FSA) or -1
CC (FSB) (see also AAR38037-38). These peptides and antibodies raised
CC against these may be used to treat or prevent spongiform encephalopathy
CC in humans, sheep or cattle. They can be used to block cellular binding
CC and aggregation of prion proteins and to stimulate the mammalian immune
CC system. These peptides may be used to distinguish between the normal
CC form of prion protein (PrPc) and the scrapie-associated form (PrPsc).
CC These peptides may include rare or synthetic amino acids or a ratio-
CC inverse peptide modification to improve resistance to enzymatic
CC degradation.
XX
XX Sequence 18 AA;
XX
XX Query Match 100.0%; Score 55; DB 14; Length 18;
XX Best Local Similarity 100.0%; Pred. No. 0.035;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 GWGQPHGG 8
XX Db 4 gwqphgg 11
XX
XX RESULT 4
XX AAR38032
XX ID AAR38032 standard; protein; 18 AA.
XX
XX AC AAR38032;
XX
XX DT 14-OCT-1993 (first entry)
XX
XX DE Human prion protein region E #2.
XX
XX KW Antigen; prion; protein; region; frame shift; repeat; mutation; PrPc;
XX FSA; FSB; subfragment; antibody; treatment; spongiform encephalopathy;
XX human; sheep; cattle; cellular binding; aggregation; mammal; scrapie;
XX immune system; PrPsc; ratio-inverse peptide; enzymatic degradation;
XX resistance.
XX
XX OS Synthetic.
XX

```

```

PH Key Location/Qualifiers
FT Misc-difference 1 /note= "One or more residues or may be absent"
FT Misc-difference 2 /note= "May be absent"
FT Misc-difference 3 /note= "May be absent"
FT Misc-difference 4 /note= "May be absent"
FT Misc-difference 5 /note= "May be absent"
FT Misc-difference 14 /note= "May be absent"
FT Misc-difference 15 /note= "May be absent"
FT Misc-difference 16 /note= "May be absent"
FT Misc-difference 17 /note= "May be absent"
FT Misc-difference 18 /note= "One or more residue or may be absent"
XX
XX WO9311155-A.
XX
XX 10-JUN-1993.
XX
XX 03-DEC-1992; 92WO-GB02246.
XX
XX 03-DEC-1991; 91GB-0025747.
XX 10-JUL-1992; 92GB-0014663.
XX
XX (PROT-) PROTEUS MOLECULAR DESIGN LTD.
XX
XX Fishleigh RV, Mee RP, Robson B;
XX WPI; 1993-196994/24.
XX
XX New polypeptide(s) contg. antigenic site of prion protein -
XX useful for treatment and diagnosis of mammalian encephalopathies
XX e.g. Creutzfeld-Jacob disease and kuru
XX
XX Claim 23; Page 71; 82pp; English.
XX
XX The sequences given in AAR38025-36 represent polypeptides derived from
CC an antigenic site, region E, of a prion protein. Prion proteins
CC comprise six regions of interest (A-F), and two related frame shift
CC peptides sequences caused by a repeating section in region E having
CC a nucleic acid coding sequence frame shift mutation of +1 (FSA) or -1
CC (FSB) (see also AAR38037-38). These peptides and antibodies raised
CC against these may be used to treat or prevent spongiform encephalopathy
CC in humans, sheep or cattle. They can be used to block cellular binding
CC and aggregation of prion proteins and to stimulate the mammalian immune
CC system. These peptides may be used to distinguish between the normal
CC form of prion protein (PrPc) and the scrapie-associated form (PrPsc).
CC These peptides may include rare or synthetic amino acids or a ratio-
CC inverse peptide modification to improve resistance to enzymatic
CC degradation.
XX
XX Sequence 18 AA;
XX
XX Query Match 100.0%; Score 55; DB 14; Length 18;
XX Best Local Similarity 100.0%; Pred. No. 0.035;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 GWGQPHGG 8
XX Db 4 gwqphgg 11
XX
XX RESULT 5
XX AAR38035
XX ID AAR38035 standard; protein; 26 AA.

```

AC AAR38035;
 DT 14-OCT-1993 (first entry)
 DE Prion protein region E #2.
 KW Antigen; prion; protein; region; frame shift; repeat; mutation; PrPc;
 KW FSA; FSB; subfragment; antibody; treatment; spongiform encephalopathy;
 KW human; sheep; cattle; cellular binding; aggregation; mammal; scrapie;
 KW immune system; PrPsc; ratio-inverso peptide; enzymatic degradation;
 KW resistance.
 XX Synthetic.
 OS W09311155-A.
 PN 10-JUN-1993.
 PD 03-DEC-1992; 92WO-GB02246.
 PF 03-DEC-1991; 91GB-0025747.
 PR 10-JUL-1992; 92GB-0014663.
 XX (PROT-) PROTEUS MOLECULAR DESIGN LTD.
 PA Fishleigh RV, Mee RP, Robson B;
 PI WPI; 1993-196994/24.
 DR New polypeptide(s) contg. antigenic site of prion protein -
 PT useful for treatment and diagnosis of mammalian encephalopathies
 PT e.g. Creutzfeld-Jacob disease and kuru
 XX Claim 24; Page 71; 82pp; English.
 XX The sequences given in AAR38025-36 represent polypeptides derived from
 CC an antigenic site, region E, of a prion protein. Prion proteins
 CC comprise six regions of interest (A-F), and two related frame shift
 CC peptides sequences caused by a repeating section in region E having
 CC a nucleic acid coding sequence frame shift mutation of +1 (FSA) or -1
 CC (FSB) (see also AAR38037-38). These peptides and antibodies raised
 CC against these may be used to treat or prevent spongiform encephalopathy
 CC in humans, sheep or cattle. They can be used to block cellular binding
 CC and aggregation of prion proteins and to stimulate the mammalian immune
 CC system. These peptides may be used to distinguish between the normal
 CC form of prion protein (PrPc) and the scrapie-associated form (PrPsc).
 CC These peptides may include rare or synthetic amino acids or a ratio-
 CC inverso peptide modification to improve resistance to enzymatic
 CC degradation.
 XX SQ Sequence 26 AA;
 Query Match 100.0%; Score 55; DB 14; Length 26;
 Best Local Similarity 100.0%; Pred. No. 0.05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GWGQPHGG 8
 DB 7 gwgqphgg 14
 RESULT 6
 AAY07999
 ID AAY07999 standard; protein; 42 AA.
 XX AC AAY07999;
 XX 08-JUL-1999 (first entry)
 XX Bovine prion protein derived peptide III.
 DE DT 08-JUL-1999 (first entry)
 DE DT Bovine prion protein derived peptide III.
 XX

KW Prion protein; PrP; human; polyclonal antiserum; immunoassay;
 KW detection; bovine; murine.
 OS Bos taurus.
 XX DE19745443-A1.
 XX 22-APR-1999.
 XX 15-OCT-1997; 97DE-1045443.
 XX 15-OCT-1997; 97DE-1045443.
 XX (HERZ/) HERZOG-MESMER A.
 XX Kiselev OI, Mesmer AH, Scheller A;
 XX WPI; 1999-255775/22.
 XX Diagnostic polyclonal antiserum specific for prion protein -
 PT obtained by immunisation with metal-containing polypeptide
 PT Disclosure; Page 4; 12pp; German.
 XX This invention describes a novel process for producing a polyclonal
 CC antiserum against a human or animal prion protein (PrP) which can be
 CC used in immunoassays for detecting PrP's. The method comprises (a)
 CC selecting a polypeptide that has a length of at least 10 amino acids and
 CC has an amino acid sequence at least 70% homologous to that of human,
 CC bovine or murine PrP in a region of at least 10 consecutive amino acids
 CC (b) binding a metal to the polypeptide by reaction with a metal compound
 CC and (c) injecting the metal-containing polypeptide into a host animal,
 CC optionally together with adjuvants, to induce production of a polyclonal
 CC antiserum.
 XX SQ Sequence 42 AA;
 Query Match 100.0%; Score 55; DB 20; Length 42;
 Best Local Similarity 100.0%; Pred. No. 0.078;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GWGQPHGG 8
 DB 4 gwgqphgg 11
 RESULT 7
 AAW70280
 ID AAW70280 standard; peptide; 178 AA.
 XX AC AAW70280;
 XX 06-NOV-1998 (first entry)
 XX Peptide sequences used to raise antibodies against prion protein.
 DE Spinal cord; cattle; sheep; pig; bovine spongiform encephalopathy;
 KW BSE; scrapie; transmissible spongiform encephalopathy; TSE;
 KW immunological assay; scrapie prion protein; prpsc.
 XX Synthetic.
 OS Key Location/Qualifiers
 XX Region 28..61
 FT /note= "Synthetic sequence used to raise antibody
 FT against prpsc"
 FT Region 90..128
 FT /note= "Synthetic sequence used to raise antibody
 FT against prpsc"
 XX W09835236-A2.
 XX

PD 13-AUG-1998.
XX
PF 06-FEB-1998; 98WO-IE00007.
XX
XX 01-MAY-1997; 97IE-0000325.
PR 06-FEB-1997; 97IE-0000081.
PR 24-MAR-1997; 97IE-0000228.
XX
PA (ENFE-) ENFER TECHNOLOGY LTD.
XX
XX O'Connor M;
PI
XX WPI; 1998-447377/38.
DR

XX Detecting pathogenic prion(s) in animal carcasses - by reaction with
XX specific labelled antibody, used to detect those carrying agents for
XX bovine spongiform encephalopathy and scrapie
XX
PS Claim 3; Page 23; 25pp; English.
XX

XX The invention claims to provide a method for detecting transmissible
XX spongiform encephalopathies (TSE) in animals and in animal carcasses.
XX The method comprises of an immunological assay whereby the animal test
XX sample is reacted with a labelled antibody against scrapie prion protein
XX (prpsc) and the amount of bound labelled antibody is then detected. The
XX anti-prpsc antibodies used in the assay are raised against fragments of
XX the present synthetic peptide shown. The peptide fragments preferred
XX by the inventors are shown in the features table. The method is claimed
XX to be useful when applied to samples, particularly a cross-section of
XX the spinal cord, from cattle, sheep and pig carcasses for detection of
XX bovine spongiform encephalopathy (BSE) or scrapie.
XX

SQ Sequence 178 AA;

Query Match 100.0%; Score 55; DB 19; Length 178;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8
| | | | | | | |
DB 61 gwqphgg 68

RESULT 8
AAB07316
ID AAB07316 standard; protein; 208 AA.

AC AAB07316;

DT 17-OCT-2000 (first entry)

DE Mouse prion protein sequence.

XX Mouse; prion protein; transmissible spongiform encephalopathy;
KW bovine spongiform encephalopathy; BSE diagnosis; TSE; PrP.

XX Mus sp.

XX Key Location/Qualifiers

FT Region 37..68
FT /note= "Repeat region consisting of tandem repeats
FT of repeat unit: PHGGGWGQ (AAB07319)"
FT Disulfide-bond 156..191
FT Modified-site 208

FT /note= "C-terminal phospho-inositol glycolipid
FT membrane anchor (-GPI)"
FT

XX WO200029850-A1.

XX 25-MAY-2000.

XX 27-OCT-1999; 99WO-FI00897.

XX 17-NOV-1998; 98FI-0002481.
PR (WALL-) WALLAC OY.
PA (BBSR-) BBSRC OFFICE.
XX
XX Hope J, Barnard GJR, Birkett CR;
XX
XX WPI; 2000-387880/33.
DR

XX Novel immunoassay for prion protein, used for the determination of
XX transmissible spongiform encephalopathies in bovines -
XX
XX Disclosure; Page 41-42; 50pp; English.

XX The present sequence is the mouse prion protein (PrP) sequence.
XX Conversion of the normal cellular form of PrP into an aggregated,
XX insoluble isoform is implicated in the pathogenesis of transmissible
XX Spongiform Encephalopathies (TSEs). Examples of TSEs include Bovine
XX Spongiform Encephalopathy (BSE), scrapie, Creutzfeldt-Jakob disease
XX (CJD) and Gerstmann-Strausler-Sheinker syndrome (GSS). The concentration
XX of this protein in body fluid or tissue samples may be measured by an
XX assay of the present invention, in which a PrP epitope is captured by an
XX antibody, which is then detected. The presence of PrP indicates TSE. PrP
XX epitopes (AAB07320-B07326) are derived from the protease resistant core
XX of PrP that is occluded when the PrP is in an aggregated state.

SQ Sequence 208 AA;

Query Match 100.0%; Score 55; DB 21; Length 208;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8
| | | | | | | |
DB 41 gwqphgg 48

RESULT 9

AAB07318
ID AAB07318 standard; protein; 208 AA.

AC AAB07318;

DT 17-OCT-2000 (first entry)

XX Human prion protein sequence.

XX Human; prion protein; transmissible spongiform encephalopathy;
KW bovine spongiform encephalopathy; BSE diagnosis; TSE; PrP.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Region 29..69
FT /note= "Repeat region consisting of tandem repeats
FT of repeat unit: PHGGGWGQ (AAB07319)"
FT Disulfide-bond 157..192
FT Modified-site 208

FT /note= "C-terminal phospho-inositol glycolipid
FT membrane anchor (-GPI)"
FT

XX WO200029850-A1.

XX 25-MAY-2000.

XX 27-OCT-1999; 99WO-FI00897.

XX 17-NOV-1998; 98FI-0002481.

XX (WALL-) WALLAC OY.

PA (BBSR-) BBSRC OFFICE.

XX PI Hope J, Barnard GJR, Birkett CR;
 XX DR WPI; 2000-387880/33.
 XX PT Novel immunoassay for prion protein, used for the determination of
 XX transmissible spongiform encephalopathies in bovines -
 XX PS Disclosure; Page 43-44; 50pp; English.
 XX CC The present sequence is the human prion protein (PrP) sequence.
 XX CC Conversion of the normal cellular form of PrP into an aggregated,
 XX CC insoluble isoform is implicated in the pathogenesis of Transmissible
 XX CC Spongiform Encephalopathies (TSEs). Examples of TSEs include Bovine
 XX CC Spongiform Encephalopathy (BSE), scrapie, Creutzfeldt-Jakob disease
 XX CC (CJD) and Gerstmann-Strausler-Sheinker syndrome (GSS). The concentration
 XX CC of this protein in body fluid or tissue samples may be measured by an
 XX CC assay of the present invention, in which a PrP epitope is captured by an
 XX CC antibody, which is then detected. The presence of PrP indicates TSE. PrP
 XX CC epitopes (AAB07320-B07326) are derived from the protease resistant core
 XX CC of PrP that is occluded when the PrP is in an aggregated state.
 XX SQ Sequence 208 AA;

Query Match 100.0%; Score 55; DB 21; Length 208;
 Best Local Similarity 100.0%; Pred. No. 0.35;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8
 Db 34 gwqphgg 41
 |||||

RESULT 10
 AAB07327
 ID AAB07327 standard; protein: 208 AA.
 XX AC AAB07327;
 XX DT 17-OCT-2000 (first entry)
 XX DE Mouse prion protein sequence.
 XX KW Mouse; prion protein; transmissible spongiform encephalopathy;
 XX KW bovine spongiform encephalopathy; TSE diagnosis; PrP.
 XX OS Mus sp.
 XX FH Key Location/Qualifiers
 XX FT Region 37..68
 FT /note= "Repeat region consisting of tandem repeats
 FT of repeat unit: PHGGGWGQ (AAB07319)"
 FT Disulfide-bond 156..191
 FT Modified-site 208
 FT /note= "C-terminal phospho-inositol glycolipid
 FT membrane anchor (-GPI)"
 XX PN WO200029849-A1.
 XX PD 25-MAY-2000.
 XX PF 27-OCT-1999; 99WO-FI00896.
 XX PR 17-NOV-1998; 98FI-0002480.
 XX PA (WALL-) WALLAC OY.
 XX PA (BBSR-) BBSRC OFFICE.
 XX PI Hope J, Barnard GJR, Birkett CR;
 XX DR WPI; 2000-395778/34.
 XX

PT New immunoassay for prion protein, used for determination of
 PT transmissible spongiform encephalopathies in mammals, comprises
 PT specific capture antibody -
 XX PS Disclosure; Page 41-42; 50pp; English.
 XX CC The present sequence is the mouse prion protein (PrP) sequence.
 XX CC Conversion of the normal cellular form of PrP into an aggregated,
 XX CC insoluble isoform is implicated in the pathogenesis of Transmissible
 XX CC Spongiform Encephalopathies (TSEs). Examples of TSEs include Bovine
 XX CC Spongiform Encephalopathy (BSE), scrapie, Creutzfeldt-Jakob disease
 XX CC (CJD) and Gerstmann-Strausler-Sheinker syndrome (GSS). The concentration
 XX CC of this protein in body fluid or tissue samples may be measured by an
 XX CC assay of the present invention, in which a PrP epitope is captured by an
 XX CC antibody, which is then detected. The presence of PrP indicates TSE. PrP
 XX CC epitopes (AAB07320-B07326) are derived from the protease resistant core
 XX CC of PrP that is occluded when the PrP is in an aggregated state.
 XX SQ Sequence 208 AA;

Query Match 100.0%; Score 55; DB 21; Length 208;
 Best Local Similarity 100.0%; Pred. No. 0.35;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8
 Db 41 gwqphgg 48
 |||||

RESULT 11
 AAB07329
 ID AAB07329 standard; protein: 208 AA.
 XX AC AAB07329;
 XX DT 17-OCT-2000 (first entry)
 XX DE Human prion protein sequence.
 XX KW Human; prion protein; transmissible spongiform encephalopathy;
 XX KW bovine spongiform encephalopathy; TSE diagnosis; PrP.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX FT Region 29..69
 FT /note= "Repeat region consisting of tandem repeats
 FT of repeat unit: PHGGGWGQ (AAB07319)"
 FT Disulfide-bond 157..192
 FT Modified-site 208
 FT /note= "C-terminal phospho-inositol glycolipid
 FT membrane anchor (-GPI)"
 XX PN WO200029849-A1.
 XX PD 25-MAY-2000.
 XX PF 27-OCT-1999; 99WO-FI00896.
 XX PR 17-NOV-1998; 98FI-0002480.
 XX PA (WALL-) WALLAC OY.
 XX PA (BBSR-) BBSRC OFFICE.
 XX PI Hope J, Barnard GJR, Birkett CR;
 XX DR WPI; 2000-399778/34.
 XX

New immunoassay for prion protein, used for determination of
 transmissible spongiform encephalopathies in mammals, comprises
 specific capture antibody -

PS Disclosure; Page 43-44; 50pp; English.

XX The present sequence is the human prion protein (PrP) sequence.
CC Conversion of the normal cellular form of PrP into an aggregated,
CC insoluble isoform is implicated in the pathogenesis of Transmissible
CC Spongiform Encephalopathies (TSEs). Examples of TSEs include Bovine
CC Spongiform Encephalopathy (BSE), scrapie, Creutzfeldt-Jakob disease
CC (CJD) and Gerstmann-Strausler-Sheinker syndrome (GSS). The concentration
CC of this protein in body fluid or tissue samples may be measured by an
CC assay of the present invention, in which a PrP epitope is captured by an
CC antibody, which is then detected. The presence of PrP indicates TSE. PrP
CC epitopes (AAB07320-B07326) are derived from the protease resistant core
CC of PrP that is occluded when the PrP is in an aggregated state.

XX SQ Sequence 208 AA;

Query Match 100.0%; Score 55; DB 21; Length 208;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGS 8
Db ||||| 34 gwqgphg 41

RESULT 12

AAB82110
ID AAB82110 standard; peptide; 208 AA.

XX AC AAB82110;

XX DT 29-JUN-2001 (first entry)

XX DE Hamster PrP peptide fragment #17.

XX Hamster; PrP: cerebroprotective; PrP conversion inhibitor; prion protein;
KW transmissible spongiform encephalopathy; TSE; neurodegenerative disease;
KW protease-sensitive prion protein; PrPsen;
XX protease-resistant prion protein; PrPres.

XX OS Cricetulus griseus.

XX PN US6211149-B1.

XX PD 03-APR-2001.

XX PF 03-AUG-1998; 98US-0128450.

XX PR 03-AUG-1998; 98US-0128450.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Chesebro BW, Caughey BW, Chabry J, Priola S;

XX WPI; 2001-315407/33.

XX New peptide comprises hamster prion protein fragment that inhibits
PT conversion of prion protein from protease-sensitive to
PT protease-resistant form, useful for diagnosis and treatment of
PT spongiform encephalopathies -

XX Disclosure; Column 29-32; 31pp; English.

XX Transmissible spongiform encephalopathies (TSE) are fatal
CC neurodegenerative diseases. These diseases are characterised by the
CC formation and accumulation, in the brain, of an abnormal proteinase K
CC resistant isoform (PrPres) of a normal protease-sensitive host-encoded
CC prion protein (PrPsen). The present invention relates to peptides
CC comprising a hamster, human or murine prion protein (PrP) fragment which
CC specifically inhibit the conversion of protease-sensitive prion protein
CC (PrPsen) to protease-resistant prion protein (PrPres). The present
CC sequence is one such peptide. The peptides of the present invention are

CC useful for diagnosis and treatment of TSE diseases.

XX SQ Sequence 208 AA;

Query Match 100.0%; Score 55; DB 22; Length 208;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGS 8
Db ||||| 34 gwqgphg 41

RESULT 13

AAB30801
ID AAB30801 standard; Protein; 211 AA.

XX AC AAB30801;

XX DT 02-APR-2001 (first entry)

XX DE Amino acid sequence of a mouse prion protein.

XX SCHAG; self-coalesce; higher-order aggregate; amyloidogenic domain;
KW aggregation; fibril; phenotypic alteration; gene therapy;
KW disease resistance; plant pigmentation; prion disease.

XX OS Mus sp.

XX PN WO200075324-A2.

XX PD 14-DEC-2000.

XX PF 09-JUN-2000; 2000WO-US15876.

XX PR 09-JUN-1999; 99US-0138833.

XX PA (ARCH-) ARCH DEV CORP.

XX PI Lindquist S, Li L, Ma J, Liu J, Sondheimer N, Scheibel T;

XX WPI; 2001-061723/07.

XX DR N-PSDB; AAC86686.

XX New nucleic acid encoding chimeric proteins with self-assembly
PT properties, useful e.g. for diagnosis and treatment of prion diseases,
PT also related aggregates, fibrils and polymers -

XX Claim 11; Page 137-138; 188pp; English.

XX The present sequence represents a prion protein. The specification
CC describes chimeric polypeptides, which comprise at least one SCHAG
CC (self-coalesce into higher-order aggregates) amino acid sequence fused
CC in frame with a polypeptide of interest (which is other than a marker
CC protein, a glutathione-S-transferase or a staphylococcal nuclear
CC protein). The specification also describes chimeric polypeptides that
CC comprises an amyloidogenic domain that causes aggregation into fibrils.
CC The chimeric polypeptides are used to prepare polymers with multiple
CC reactivities, e.g. derivatised with enzymes, or specific binding
CC partners, and useful e.g. for performing multi-step chemical reactions.
CC They can be used to create an inducible, or stable phenotypic alteration
CC in a cell, e.g. for gene therapy, protein production, imparting disease
CC resistance to plants, altering plant pigmentation and for diagnosis
CC and treatment of prion diseases.

XX SQ Sequence 211 AA;

Query Match 100.0%; Score 55; DB 22; Length 211;
Best Local Similarity 100.0%; Pred. No. 0.36;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8
 |||||
 Db 43 gwqphgg 50

RESULT 14

AAB30802
 ID AAB30802 standard; Protein; 212 AA.

XX AC AAB30802;

XX DT 02-APR-2001 (first entry)

XX DE Amino acid sequence of a Syrian hamster prion protein.

XX KW SCHAG: self-coalesce; higher-order aggregate; amyloidogenic domain;
 aggregation; fibril; phenotypic alteration; gene therapy;
 XX KW disease resistance; plant pigmentation; prion disease.

XX OS Mesocricetus auratus.

XX PN WO200075324-A2.

XX PD 14-DEC-2000.

XX PF 09-JUN-2000; 2000WO-US15876.

XX PR 09-JUN-1999; 99US-0138833.

XX PA (ARCH-) ARCH DEV CORP.

XX PI Lindquist S, Li L, Ma J, Liu J, Sondheimer N, Scheibel T;

XX DR WPI; 2001-061723/07.

XX DR N-PSDB; RAC86687.

XX FT New nucleic acid encoding chimeric proteins with self-assembly
 properties, useful e.g. for diagnosis and treatment of prion diseases,
 XX FT also related aggregates, fibrils and polymers.

XX PS Claim 11; Page 139-140; 188pp; English.

XX CC The present sequence represents a prion protein. The specification
 CC describes chimeric polypeptides, which comprise at least one SCHAG.
 CC (self-coalesces into higher-order aggregates) amino acid sequence fused
 CC in frame with a polypeptide of interest (which is other than a marker
 CC protein, a glutathione-S-transferase or a staphylococcal nuclear
 CC protein). The specification also describes chimeric polypeptides that
 CC comprises an amyloidogenic domain that causes aggregation into fibrils.
 CC The chimeric polypeptides are used to prepare polymers with multiple
 CC reactivities, e.g. derivatised with enzymes, or specific binding
 CC partners, and useful e.g. for performing multi-step chemical reactions.
 CC They can be used to create an inducible, or stable phenotypic alteration
 CC in a cell, e.g. for gene therapy, protein production, imparting disease
 CC resistance to plants, altering plant pigmentation and for diagnosis
 CC and treatment of prion diseases.

XX SQ Sequence 212 AA;

Query Match 100.0%; Score 55; DB 22; Length 212;
 Best Local Similarity 100.0%; Pred. No. 0.36;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8
 |||||
 Db 44 gwqphgg 51

RESULT 15

AAB07317
 ID AAB07317 standard; protein; 217 AA.

XX

AC AAB07317;

XX DT 17-OCT-2000 (first entry)

XX DE Cattle prion protein sequence.

XX KW Cattle; prion protein; transmissible spongiform encephalopathy;
 XX KW bovine spongiform encephalopathy; BSE diagnosis; TSE; PrP.

XX OS Bos bovis.

XX FH Key Location/Qualifiers

XX FT Region 37..79
 FT /note= "Repeat region consisting of tandem repeats
 FT of repeat unit: PHGGWGQ (AAB07319)";

XX FT Disulfide-bond 166..201

XX FT Modified-site 217

XX FT /note= "C-terminal phospho-inositol glycolipid
 XX FT membrane anchor (-GPI)";

XX PN WO200029850-A1.

XX PD 25-MAY-2000.

XX PF 27-OCT-1999; 99WO-FI00897.

XX PR 17-NOV-1998; 98FI-0002481.

XX XX (WALL-) WALLAC OY.

XX PA (BBSR-) BBSRC OFFICE.

XX XX Hope J, Barnard GJR, Birkett CR;

XX XX WPI; 2000-387880/33.

XX FT Novel immunoassay for prion protein, used for the determination of
 XX FT transmissible spongiform encephalopathies in bovines -

XX PS Disclosure; Page 42-43; 50pp; English.

XX CC The present sequence is the cattle prion protein (PrP) sequence.
 CC Conversion of the normal cellular form of PrP into an aggregated,
 CC insoluble isoform is implicated in the pathogenesis of transmissible
 CC Spongiform Encephalopathies (TSEs). Examples of TSEs include Bovine
 CC Spongiform Encephalopathy (BSE), scrapie, Creutzfeldt-Jakob disease
 CC (CJD) and Gerstmann-Strausler-Sheinker syndrome (GSS). The concentration
 CC of this protein in body fluid or tissue samples may be measured by an
 CC assay of the present invention, in which a PrP epitope is captured by an
 CC antibody, which is then detected. The presence of PrP indicates BSE. PrP
 CC epitopes (AAB07320-807326) are derived from the protease resistant core
 CC of PrP that is occluded when the PrP is in an aggregated state.

XX SQ Sequence 217 AA;

Query Match 100.0%; Score 55; DB 21; Length 217;
 Best Local Similarity 100.0%; Pred. No. 0.37;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8
 |||||
 Db 35 gwqphgg 42

Search completed: August 6, 2002, 10:38:47
 Job time: 124 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 10:36:43 ; Search time 26.75 Seconds
(without alignments)
7.305 Million cell updates/sec

Title: US-09-543-188A-1

Perfect score: 55

Sequence: 1 GWGQPHGG 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	18	1	US-08-244-701B-24
2	55	100.0	18	1	US-08-244-701B-27
3	55	100.0	18	1	US-08-244-701B-30
4	55	100.0	26	1	US-08-244-701B-46
5	55	100.0	208	4	US-09-128-450-18
6	55	100.0	208	4	US-09-823-494-18
7	55	100.0	253	1	US-08-242-188-2
8	55	100.0	253	1	US-08-509-261A-2
9	55	100.0	253	1	US-08-660-626-8
10	55	100.0	253	1	US-08-692-892-2
11	55	100.0	253	2	US-08-713-939A-2
12	55	100.0	253	2	US-08-868-162A-22
13	55	100.0	253	4	US-09-031-168-8
14	55	100.0	253	4	US-09-128-450-20
15	55	100.0	253	4	US-09-036-579-2
16	55	100.0	253	4	US-09-823-494-20
17	55	100.0	254	1	US-08-242-188-1
18	55	100.0	254	1	US-08-509-261A-1
19	55	100.0	254	1	US-08-660-626-7
20	55	100.0	254	1	US-08-692-892-1
21	55	100.0	254	2	US-08-713-939A-1
22	55	100.0	254	2	US-08-868-162A-21
23	55	100.0	254	4	US-09-031-168-7
24	55	100.0	254	4	US-09-128-450-19
25	55	100.0	254	4	US-09-128-450-26
26	55	100.0	254	4	US-09-128-450-28
27	55	100.0	254	4	US-09-036-579-1

Sequence 19, Appl
Sequence 26, Appl
Sequence 28, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 10, Appl
Sequence 4, Appl
Sequence 24, Appl
Sequence 10, Appl
Sequence 22, Appl
Sequence 22, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 9, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 3, Appl

28 55 100.0 254 4 US-09-823-494-19
29 55 100.0 254 4 US-09-823-494-26
30 55 100.0 254 4 US-09-823-494-28
31 55 100.0 255 1 US-08-242-188-4
32 55 100.0 255 1 US-08-509-261A-4
33 55 100.0 255 1 US-08-660-626-10
34 55 100.0 255 1 US-08-692-892-4
35 55 100.0 255 2 US-08-713-939A-4
36 55 100.0 255 2 US-08-868-162A-24
37 55 100.0 255 4 US-09-031-168-10
38 55 100.0 255 4 US-09-036-579-4
39 55 100.0 256 4 US-09-128-450-22
40 55 100.0 256 4 US-09-823-494-22
41 55 100.0 263 1 US-08-242-188-3
42 55 100.0 263 1 US-08-509-261A-3
43 55 100.0 263 1 US-08-660-626-9
44 55 100.0 263 1 US-08-692-892-3
45 55 100.0 263 2 US-08-713-939A-3

ALIGNMENTS

RESULT 1
US-08-244-701B-24
; Sequence 24, Application US/08244701B
; Patent No. 5773572
; GENERAL INFORMATION:
; APPLICANT: Fishleigh, Robert V.
; APPLICANT: Robson, Barry
; TITLE OF INVENTION: Fragments of Prion Proteins
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/244,701B
; FILING DATE: 02-JUN-1994
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Fanucci, Allan A.
; REGISTRATION NUMBER: 30,256
; REFERENCE/DOCKET NUMBER: 8080-007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /label= X
; OTHER INFORMATION: /note= "X may be absent or present independently
; OF Y and denotes one or more amino acid(s)"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 18

OTHER INFORMATION: /label- Y
OTHER INFORMATION: /note- "Y may be absent or present independently
OTHER INFORMATION: of X and denotes one or more amino acid(s)"
US-08-244-701B-24

Query Match 100.0%; Score 55; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWQPHGG 8
Db 4 GWQPHGG 11

RESULT 2
US-08-244-701B-27
; Sequence 27, Application US/08244701B
; Patent No. 5773572
; GENERAL INFORMATION:
; APPLICANT: Fishleigh, Robert V.
; APPLICANT: Robson, Barry
; APPLICANT: Mee, Roger P.
; TITLE OF INVENTION: Fragments of Prion Proteins
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/244,701B
FILING DATE: 02-JUN-1994
CLASSIFICATION: 436

ATTORNEY/AGENT INFORMATION:
NAME: Panucci, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 8080-007
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:

NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /label- X
OTHER INFORMATION: /note- "X may be absent or present independently
OTHER INFORMATION: of Y and denotes one or more amino acid(s)"
FEATURE:

NAME/KEY: Modified-site
LOCATION: 18
OTHER INFORMATION: /label- Y
OTHER INFORMATION: /note- "Y may be absent or present independently
OTHER INFORMATION: of X and denotes one or more amino acid(s)"
US-08-244-701B-27

Query Match 100.0%; Score 55; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.015;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GWQPHGG 8
Db 4 GWQPHGG 11

RESULT 3
US-08-244-701B-30
; Sequence 30, Application US/08244701B
; Patent No. 5773572
; GENERAL INFORMATION:
; APPLICANT: Fishleigh, Robert V.
; APPLICANT: Robson, Barry
; APPLICANT: Mee, Roger P.
; TITLE OF INVENTION: Fragments of Prion Proteins
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/244,701B
FILING DATE: 02-JUN-1994
CLASSIFICATION: 436

ATTORNEY/AGENT INFORMATION:
NAME: Panucci, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 8080-007
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:

NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /label- X
OTHER INFORMATION: /note- "X is absent or present independently of Y
OTHER INFORMATION: and denotes one or more amino acid(s)"
FEATURE:

NAME/KEY: Modified-site
LOCATION: 18
OTHER INFORMATION: /label- Y
OTHER INFORMATION: /note- "Y may be absent or present independently
OTHER INFORMATION: of X and denotes one or more amino acid(s)"
US-08-244-701B-30

Query Match 100.0%; Score 55; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWQPHGG 8
Db 4 GWQPHGG 11

RESULT 4

US-08-244-701B-46
 ; Sequence 46, Application US/08244701B
 ; Patent No. 5773572
 ; GENERAL INFORMATION:
 ; APPLICANT: Fishleigh, Robert V.
 ; APPLICANT: Robson, Barry
 ; APPLICANT: Mee, Roger P.
 ; TITLE OF INVENTION: Fragments of Prion Proteins
 ; NUMBER OF SEQUENCES: 67
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/244,701B
 ; FILING DATE: 02-JUN-1994
 ; CLASSIFICATION: 436
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fanucci, Allan A.
 ; REGISTRATION NUMBER: 30,256
 ; REFERENCE/DOCKET NUMBER: 8080-007
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-8864/9741
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 46:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 26 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-244-701B-46

Query Match 100.0%; Score 55; DB 1; Length 26;
 Best Local Similarity 100.0%; Pred. No. 0.021;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8
 Db 7 GWGQPHGG 14

RESULT 5
 US-09-128-450-18
 ; Sequence 18, Application US/09128450
 ; Patent No. 6211149
 ; GENERAL INFORMATION:
 ; APPLICANT: Chesebro, Bruce W
 ; APPLICANT: Caughey, Byron W
 ; APPLICANT: Caughey, Joelle
 ; APPLICANT: Priola, Susette
 ; TITLE OF INVENTION: Inhibitors of Formation of Protease Resistant Prion
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Karl Bosicevic
 ; STREET: 2200 Sand Hill Road
 ; CITY: Menlo Park
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94025
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS

; NAME/KEY: NON_TER
 ; LOCATION: (1)..(2)
 ; US-09-128-450-18

Query Match 100.0%; Score 55; DB 4; Length 208;
 Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8
 Db 34 GWGQPHGG 41

RESULT 6
 US-09-823-494-18
 ; Sequence 18, Application US/09823494
 ; Patent No. 6355610
 ; GENERAL INFORMATION:
 ; APPLICANT: Chesebro, Bruce W
 ; APPLICANT: Caughey, Byron W
 ; APPLICANT: Caughey, Joelle
 ; APPLICANT: Priola, Susette
 ; TITLE OF INVENTION: Inhibitors of Formation of Protease Resistant Prion
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Karl Bosicevic
 ; STREET: 2200 Sand Hill Road
 ; CITY: Menlo Park
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94025
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/823,494
 ; FILING DATE: 02-JUN-1994
 ; CLASSIFICATION: 436
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fanucci, Allan A.
 ; REGISTRATION NUMBER: 30,256
 ; REFERENCE/DOCKET NUMBER: 8080-007
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-8864/9741
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 46:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 26 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-09-823-494-18

Query Match 100.0%; Score 55; DB 4; Length 208;
 Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8
 Db 34 GWGQPHGG 41

RESULT 7
 US-08-242-188-2
 ; Sequence 2, Application US/08242188
 ; Patent No. 5565186
 ; GENERAL INFORMATION:
 ; APPLICANT: Prusiner, Stanley B.
 ; APPLICANT: Scott, Michael R.
 ; APPLICANT: Telling, Glenn
 ; TITLE OF INVENTION: METHOD OF DETECTING PRIONS IN A SAMPLE
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Karl Bosicevic
 ; STREET: 2200 Sand Hill Road
 ; CITY: Menlo Park
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94025
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/242,188
FILING DATE: 13-MAY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Bosicevic, Karl
REGISTRATION NUMBER: 28,807
REFERENCE/DOCKET NUMBER: 06510/014001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 854-5277
TELEFAX: (415) 854-0875
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: HUMAN PRION PROTEIN, HuPrp
US-08-242-188-2

Query Match 100.0%; Score 55; DB 1; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8
|||||||
Db 56 GWGQPHGG 63

RESULT 8
US-08-509-261A-2
Sequence 2, Application US/08509261A
Patent No. 5763244
GENERAL INFORMATION:
APPLICANT: Prusiner, Stanley B.
APPLICANT: Scott, Michael R.
APPLICANT: Telling, Glenn
TITLE OF INVENTION: Method of Detecting Prions
TITLE OF INVENTION: In a Sample and Transgenic Animal Used fore
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bozicevic & Reed, LLP
STREET: 285 Hamilton Avenue, Suite 200
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/509,261A
FILING DATE: 31-JUL-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Bozicevic, Karl
REGISTRATION NUMBER: 28,807
REFERENCE/DOCKET NUMBER: 6510-030001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3400
TELEFAX: 650 327-3231
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

LENGTH: 253 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-509-261A-2

Query Match 100.0%; Score 55; DB 1; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8
|||||||
Db 56 GWGQPHGG 63

RESULT 9
US-08-660-626-8
Sequence 8, Application US/08660626
Patent No. 5789655
GENERAL INFORMATION:
APPLICANT: Stanley B. Prusiner
APPLICANT: Glenn C. Telling
APPLICANT: Fred E. Cohen
APPLICANT: Michael R. Scott
TITLE OF INVENTION: TRANSGENIC ANIMALS EXPRESSING
TITLE OF INVENTION: EPITOPE-TAGGED PROTEINS
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 2200 Sand Hill Road, Suite 100
CITY: Menlo Park
STATE: California
COUNTRY: USA
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/660,626
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Valeta Greg
REGISTRATION NUMBER: 35,127
REFERENCE/DOCKET NUMBER: 07532/003001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 322-5070
TELEFAX: (415) 854-0875
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: HUMAN PRION PROTEIN, HuPrp
US-08-660-626-8

Query Match 100.0%; Score 55; DB 1; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8
|||||||
Db 56 GWGQPHGG 63

RESULT 10

US-08-692-892-2
; Sequence 2, Application US/08692892
; Patent No. 5792901
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley B.
; APPLICANT: Scott, Michael R.
; APPLICANT: Telling, Glenn
; TITLE OF INVENTION: DETECTING PRIONS IN A SAMPLE AND
; TITLE OF INVENTION: DETECTING PRIONS IN A SAMPLE AND
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Karl Bozicevic
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/692,892
; FILING DATE: 30-JULY-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bozicevic, Karl
; REGISTRATION NUMBER: 28,807
; REFERENCE/DOCKET NUMBER: 06510/056001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 322-5070
; TELEFAX: (415) 854-0875
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: HUMAN PRION PROTEIN, HuPrP
US-08-692-892-2

Query Match 100.0%; Score 55; DB 1; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8
|||||
Db 56 GWGQPHGG 63

RESULT 11
US-08-713-939A-2
; Sequence 2, Application US/08713939A
; Patent No. 5846533
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley B.
; APPLICANT: Williamson, R. Anthony
; APPLICANT: Burton, Dennis R.
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR NATIVE PrP
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/713,939A
; FILING DATE: 13-SEP-1996
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bozicevic, Karl
; REGISTRATION NUMBER: 28,807
; REFERENCE/DOCKET NUMBER: 06510/059001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-5277
; TELEFAX: 415-854-0875
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-713-939A-2

Query Match 100.0%; Score 55; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8
|||||
Db 56 GWGQPHGG 63

RESULT 12
US-08-868-162A-22
; Sequence 22, Application US/08868162A
; Patent No. 5962669
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley
; APPLICANT: Cohen, Fred
; APPLICANT: James, Thomas
; APPLICANT: Kaneko, Kiyotoshi
; TITLE OF INVENTION: Prion Protein Modulator Factor
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/868,162A
; FILING DATE: 03-JUN-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bozicevic, Karl
; REGISTRATION NUMBER: 28,807
; REFERENCE/DOCKET NUMBER: 6510-083001
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-327-3400
; TELEFAX: 650 327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: HUMAN PRION PROTEIN, HuPrP
; US-08-868-162A-22

Query Match 100.0%; Score 55; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8
Db 56 GWGQPHGG 63

RESULT 13

US-09-031-168-8
; Sequence 8, Application US/09031168
; Patent No. 6150583

; GENERAL INFORMATION:
; APPLICANT: Stanley B. Prusiner
; APPLICANT: Glenn C. Telling
; APPLICANT: Fred E. Cohen
; APPLICANT: Michael R. Scott
; TITLE OF INVENTION: TRANSGENIC ANIMALS EXPRESSING
; TITLE OF INVENTION: EPIOTOPE-TAGGED PROTEINS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031.168
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/660,626
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Valeta Gregg
; REGISTRATION NUMBER: 35,127
; REFERENCE/DOCKET NUMBER: 07537/003001

; TELEPHONE: (415) 322-5070
; TELEFAX: (415) 854-0875
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: HUMAN PRION PROTEIN, HuPrP
; US-09-031-168-8

Query Match 100.0%; Score 55; DB 4; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8
Db 56 GWGQPHGG 63

RESULT 14

US-09-128-450-20
; Sequence 20, Application US/09128450
; Patent No. 6211149

; GENERAL INFORMATION:
; APPLICANT: Chesebro, Bruce W
; APPLICANT: Caughey, Byron W
; APPLICANT: Chabry, Joelle
; APPLICANT: Priola, Susette
; TITLE OF INVENTION: Inhibitors of Formation of Protease Resistant Prion
; FILE REFERENCE: 50121
; CURRENT APPLICATION NUMBER: US/09/128,450
; CURRENT FILING DATE: 1998-08-03
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-128-450-20

Query Match 100.0%; Score 55; DB 4; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8
Db 56 GWGQPHGG 63

RESULT 15

US-09-036-579-2
; Sequence 2, Application US/09036579
; Patent No. 6290954

; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley B.
; APPLICANT: Williamson, R. Anthony
; APPLICANT: Burton, Dennis R.
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR NATIVE PrP
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94025

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,579
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/713,939
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bozicevic, Karl

REGISTRATION NUMBER: 28,807
REFERENCE/DOCKET NUMBER: 06510/059001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-854-5277
TELEFAX: 415-854-0875
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-036-579-2

Query Match 100.0%; Score 55; DB 4; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GWGQPHGG 8
| | | | | | | |
Db 56 GWGQPHGG 63

Search completed: August 6, 2002, 10:39:20
Job time: 157 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 10:36:43 ; Search time 35.27 Seconds
(without alignments)
21.795 Million cell updates/sec

Title: US-09-543-188A-1

Perfect score: 55

Sequence: 1 GWGQPHGG 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	226	2 A53892	prion-related prot
2	55	100.0	232	2 S71041	major prion protei
3	55	100.0	239	2 S53633	major prion protei
4	55	100.0	241	2 S71056	major prion protei
5	55	100.0	241	2 S71048	major prion protei
6	55	100.0	245	2 S53627	major prion protei
7	55	100.0	245	2 S71045	major prion protei
8	55	100.0	252	2 I61848	major prion protei
9	55	100.0	252	2 S53634	major prion protei
10	55	100.0	252	2 S53631	major prion protei
11	55	100.0	252	2 JC6175	prion protei - ra
12	55	100.0	253	1 UJHU	major prion protei
13	55	100.0	253	2 I37032	major prion protei
14	55	100.0	253	2 I61847	major prion protei
15	55	100.0	253	2 S53635	prion protei - si
16	55	100.0	253	2 I84423	major prion protei
17	55	100.0	253	2 S53618	major prion protei
18	55	100.0	253	2 S53619	major prion protei
19	55	100.0	253	2 S53620	major prion protei
20	55	100.0	253	2 S71055	major prion protei
21	55	100.0	253	2 S53623	major prion protei
22	55	100.0	253	2 S53624	major prion protei
23	55	100.0	253	2 S53625	major prion protei
24	55	100.0	253	2 S53617	major prion protei
25	55	100.0	253	2 S53614	major prion protei
26	55	100.0	253	2 S53616	major prion protei
27	55	100.0	254	1 UJHVIH	major prion PrP-Sc
28	55	100.0	254	2 A34759	prion protei - Ch
29	55	100.0	254	2 B34759	prion protei - go

30	55	100.0	254	2 A23544	major prion protei
31	55	100.0	256	2 JU0268	major prion protei
32	55	100.0	256	2 S37149	prion protein - go
33	55	100.0	256	2 A54281	major prion protei
34	55	100.0	257	2 A23545	major prion PrP27-
35	55	100.0	257	2 JQ1900	major prion protei
36	55	100.0	260	2 S53629	major prion protei
37	55	100.0	264	2 S37137	prion protein - gr
38	55	100.0	264	2 A54330	major prion protei
39	43	78.2	381	2 A44323	pentaxin PrX3 prec
40	41	74.5	395	2 T08350	hypothetical prote
41	41	74.5	521	2 E64862	probable membrane
42	40	72.7	327	2 E87218	probable prephenat
43	40	72.7	1161	2 T45294	hypothetical prote
44	39	70.9	69	2 B89016	protein B0213.5 [i
45	39	70.9	120	2 F86824	hypothetical prote

ALIGNMENTS

RESULT 1

A53892

prion-related protein - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 13-Aug-1999

C:Accession: A53892

R:Liao, Y.C.; Tokes, Z.; Llm, E.; Lackey, A.; Woo, C.H.; Button, J.D.; Clawson, G.A.

Lab. Invest. 57, 370-374, 1987

A:Title: Cloning of rat "prion-related protein" cDNA.

A:Reference number: A53892

A:Accession: A53892

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-226 <LIA>

A:Cross-references: GS:M20313; NID:g206391; PIDN:AAA41947.1; PID:g206392

C:Superfamily: major prion protein

Query Match 100.0%; Score 55; DB 2; Length 226;

Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8

|||||||

Db 36 GWGQPHGG 43

RESULT 2

S71041

major prion protein - black-handed spider monkey (fragment)

C:Species: Ateles geoffroyi (black-handed spider monkey)

C>Date: 27-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 13-Aug-1999

C:Accession: S71041; S53630

R:Schatz, H.M.

submitted to the EMBL Data Library, April 1994

A:Reference number: S71041

A:Accession: S71041

A:Molecule type: DNA

A:Residues: 1-232 <SCH>

A:Cross-references: EMBL:U08309; NID:g474376; PIDN:AAC50097.1; PID:g474377

R:Schatz, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.

J. Mol. Biol. 245, 362-374, 1995

A:Title: Prion protein gene variation among primates.

A:Reference number: S53614; MUID:95139066

A:Accession: S53630

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-194, 'R', 196-231 <SCW>

A:Cross-references: EMBL:U08309

C:Superfamily: major prion protein

C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane

Query Match 100.0%; Score 55; DB 2; Length 232;
Best Local Similarity 100.0%; Pred. No. 0.07;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8
|||||
Db 48 GWGQPHGG 55

RESULT 3

major prion protein - douroucouli (fragment)
C:Species: Aotus trivirgatus (douroucouli, night monkey, owl monkey)
C:Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 13-Aug-1999
C:Accession: S53633; S71042
R:Schaezel, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.
J. Mol. Biol. 245, 362-374, 1995
A:Title: Prion protein gene variation among primates.
A:Reference number: S53614; MUID:95139066
A:Accession: S53633
A:Molecule type: DNA
A:Residues: 1-239 <SCH>
R:Schaezel, H.M.
A:Reference number: S71041
A:Accession: S71042
A:Molecule type: DNA
A:Residues: 1-202, 'E', 204-239 <SCW>
A:Cross-references: EMBL:U08293; NID:9474344; PIDN:AAC50082.1; PID:9474345
C:Superfamily: major prion protein
C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane protein

Query Match 100.0%; Score 55; DB 2; Length 239;
Best Local Similarity 100.0%; Pred. No. 0.07;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8
|||||
Db 48 GWGQPHGG 55

RESULT 4

major prion protein - mandrill (fragment)
C:Species: Papio sphinx, Mandrillus sphinx (mandrill)
C:Date: 27-Oct-1996 #sequence_revision 14-Feb-1997 #text_change 13-Aug-1999
C:Accession: S71056; S53621
R:Schaezel, H.M.
submitted to the EMBL Data Library, April 1994
A:Reference number: S71041
A:Accession: S71056
A:Molecule type: DNA
A:Residues: 1-241 <SCH>
A:Cross-references: EMBL:U08303; NID:9474364; PIDN:AAC50091.1; PID:9474365
R:Schaezel, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.
J. Mol. Biol. 245, 362-374, 1995
A:Title: Prion protein gene variation among primates.
A:Reference number: S53614; MUID:95139066
A:Accession: S53621
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-203, 'R', 205-240 <SCW>
A:Cross-references: EMBL:U08303
C:Superfamily: major prion protein
C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane protein

Query Match 100.0%; Score 55; DB 2; Length 241;
Best Local Similarity 100.0%; Pred. No. 0.073;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8
|||||
Db 49 GWGQPHGG 56

RESULT 5

S71048
major prion protein - Callicebus moloch (fragment)
C:Species: Callicebus moloch
C:Date: 27-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 13-Aug-1999
C:Accession: S71048; S53632
R:Schaezel, H.M.
submitted to the EMBL Data Library, April 1994
A:Reference number: S71041
A:Accession: S71048
A:Molecule type: DNA
A:Residues: 1-241 <SCH>
A:Cross-references: EMBL:U08312; NID:9475585; PIDN:AAC50100.1; PID:9475586
R:Schaezel, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.
J. Mol. Biol. 245, 362-374, 1995
A:Title: Prion protein gene variation among primates.
A:Reference number: S53614; MUID:95139066
A:Accession: S53632
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-203, 'R', 205-240 <SCW>
A:Cross-references: EMBL:U08312
C:Superfamily: major prion protein
C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane

Query Match 100.0%; Score 55; DB 2; Length 241;
Best Local Similarity 100.0%; Pred. No. 0.073;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8
|||||
Db 57 GWGQPHGG 64

RESULT 6

S53627
major prion protein - green monkey
C:Species: Cercopithecus aethiops (green monkey, grivet)
C:Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 13-Aug-1999
C:Accession: S53627; S71043
R:Schaezel, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.
J. Mol. Biol. 245, 362-374, 1995
A:Title: Prion protein gene variation among primates.
A:Reference number: S53614; MUID:95139066
A:Accession: S53627
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-245 <SCH>
A:Cross-references: EMBL:U08291
R:Schaezel, H.M.
submitted to the EMBL Data Library, April 1994
A:Reference number: S71041
A:Accession: S71043
A:Molecule type: DNA
A:Residues: 1-10, 'V', 12-202, 'E', 204-245 <SCW>
A:Cross-references: EMBL:U08291; NID:9474340; PIDN:AAC50080.1; PID:9474341
C:Superfamily: major prion protein
C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane

Query Match 100.0%; Score 55; DB 2; Length 245;
Best Local Similarity 100.0%; Pred. No. 0.074;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8

Db 56 GWGQPHGG 63
|||||

RESULT 7

S71045

major prion protein - Cercopithecus diana

C;Species: Cercopithecus diana

C;Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 13-Aug-1999

C;Accession: S71045; S53628

R;Schatzl, H.M.

submitted to the EMBL Data Library, April 1994

A;Reference number: S71041

A;Accession: S71045

A;Molecule type: DNA

A;Residues: 1-245 <SCH>

A;Cross-references: EMBL:U08292; NID:g474342; PIDN:AAAC50081.1; PID:g474343

R;Schatzl, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.

J. Mol. Biol. 245, 362-374, 1995

A;Title: Prion protein gene variation among primates.

A;Reference number: S53614; MUID:95139066

A;Accession: S53628

A;Status: nucleic acid sequence not shown

A;Molecule type: DNA

A;Residues: 8-10, 'L', 12-202, 'R', 204-239 <SCW>

A;Cross-references: EMBL:U08292

C;Superfamily: major prion protein

C;Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane protein

Query Match

Best Local Similarity 100.0%; Score 55; DB 2; Length 245;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8

|||||

Db 56 GWGQPHGG 63

RESULT 8

161848

major prion protein precursor - common squirrel monkey

C;Species: Saimiri sciureus (common squirrel monkey)

C;Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 13-Aug-1999

C;Accession: I61848

R;Cervenakova, L.; Brown, P.; Goldfarb, L.G.; Nagle, J.; Pettrone, K.; Rubenstein, R.;

Proc. Natl. Acad. Sci. U.S.A. 91, 12159-12162, 1994

A;Title: Infectious amyloid precursor gene sequences in primates used for experimental

A;Reference number: I36907; MUID:95083661

A;Accession: I61848

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-252 <RES>

A;Cross-references: EMBL:U15165; NID:g595852; PIDN:AAA68636.1; PID:g595853

C;Superfamily: major prion protein

Query Match

Best Local Similarity 100.0%; Score 55; DB 2; Length 252;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8

|||||

Db 55 GWGQPHGG 62

RESULT 9

S53634

major prion protein - common marmoset

C;Species: Callithrix jacchus (common marmoset)

C;Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 13-Aug-1999

C;Accession: S53634; S71047

R;Schatzl, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.

J. Mol. Biol. 245, 362-374, 1995
A;Title: Prion protein gene variation among primates.

A;Reference number: S53614; MUID:95139066

A;Accession: S53634

A;Status: nucleic acid sequence not shown

A;Molecule type: DNA

A;Residues: 1-252 <SCH>

A;Cross-references: EMBL:U08304

R;Schatzl, H.M.

submitted to the EMBL Data Library, April 1994

A;Reference number: S71041

A;Accession: S71047

A;Molecule type: DNA

A;Residues: 1-209, 'E', 211-252 <SCW>

A;Cross-references: EMBL:U08304; NID:g474366; PIDN:AAAC50092.1; PID:g474367

C;Superfamily: major prion protein

C;Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane

Query Match

Best Local Similarity 100.0%; Score 55; DB 2; Length 252;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8

|||||

Db 55 GWGQPHGG 62

RESULT 10

S53631

major prion protein - brown capuchin

C;Species: Cebus apella (brown capuchin, black-capped capuchin)

C;Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 13-Aug-1999

C;Accession: S53631; S71044

R;Schatzl, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.

J. Mol. Biol. 245, 362-374, 1995

A;Title: Prion protein gene variation among primates.

A;Reference number: S53614; MUID:95139066

A;Accession: S53631

A;Status: nucleic acid sequence not shown

A;Molecule type: DNA

A;Residues: 1-252 <SCH>

A;Cross-references: EMBL:U08295

R;Schatzl, H.M.

submitted to the EMBL Data Library, April 1994

A;Reference number: S71041

A;Accession: S71044

A;Molecule type: DNA

A;Residues: 1-209, 'E', 211-252 <SCW>

A;Cross-references: EMBL:U08295; NID:g474348; PIDN:AAAC50084.1; PID:g474349

C;Superfamily: major prion protein

C;Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane

Query Match

Best Local Similarity 100.0%; Score 55; DB 2; Length 252;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8

|||||

Db 55 GWGQPHGG 62

RESULT 11

JC6175

prion protein - rabbit

C;Species: Oryctolagus cuniculus (domestic rabbit)

C;Date: 11-Apr-1997 #sequence_revision 09-May-1997 #text_change 13-Aug-1999

C;Accession: JC6175

R;Loftus, B.; Rogers, M.

Gene 184, 215-219, 1997

A;Title: Characterization of a prion protein (PrP) gene from rabbit; a species with a

A;Reference number: JC6175; MUID:97183665

Search completed: August 6, 2002, 10:37:27
Job time: 44 sec

```
RESULT 14
161847
major prion protein precursor - chimpanzee
C:Species: Pan troglodytes (chimpanzee)
C>Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 13-Aug-1999
C:Accession: I61847; S71060; S53615
R:Carvenakova, L.; Brown, P.; Goldfarb, L.G.; Nagle, J.; Pettrone, K.; Rubenstein, R.;
Proc. Natl. Acad. Sci. U.S.A. 91, 12159-12162, 1994
A:Title: Infectious amyloid precursor gene sequences in primates used for experimental
A:Reference number: I36907; MUID:95083661
A:Accession: I61847
A>Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-253 <RES>
A:Cross-references: EMBL:U15039; NID:g609303; PIDN:AAA68632.1; PID:g609304
R:Schatzl, H.M.
submitted to the EMBL Data Library, April 1994
A:Reference number: S71041
A:Accession: S71060
A:Molecule type: DNA
A:Residues: 1-253 <SCW>
A:Cross-references: EMBL:U08296; NID:g474350; PIDN:AAC50085.1; PID:g474351
R:Schatzl, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.
J. Mol. Biol. 245, 362-374, 1995
A:Title: Prion protein gene variation among primates.
A:Reference number: S53614; MUID:95139066
A:Accession: S53615
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-210, 'R', 212-253 <SCH>
A:Cross-references: EMBL:U08296
C:Superfamily: major prion protein
C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane pro

Query Match 100.0%; Score 55; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.076;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GWGQPHGG 8
DB 56 GWGQPHGG 63

RESULT 15
S53635
prion protein - siamang
C:Species: Hylobates syndactylus (siamang)
C>Date: 15-Jul-1995 #sequence_revision 19-Apr-1996 #text_change 13-Aug-1999
C:Accession: S53635
R:Schatzl, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.
J. Mol. Biol. 245, 362-374, 1995
A:Title: Prion protein gene variation among primates.
A:Reference number: S53614; MUID:95139066
A:Accession: S53635
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-253 <SCH>
A:Cross-references: EMBL:U08308; NID:g474374; PIDN:AAC50096.1; PID:g474375
A>Note: the source was designated as Sympalangus syndactylus
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1994
C:Superfamily: major prion protein

Query Match 100.0%; Score 55; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.076;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GWGQPHGG 8
DB 56 GWGQPHGG 63
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 10:36:43 ; Search time 18.57 Seconds
(without alignments)
16.680 Million cell updates/sec

Title: US-09-543-188A-1

Perfect score: 55

Sequence: 1 CWGQPHGG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	55	100.0	232	1	PRIO_ATEGE	P40246	ateles geof
2	55	100.0	238	1	PRIO_CERAT	Q95145	cercopithec
3	55	100.0	238	1	PRIO_ATEGE	Q95270	theropithec
4	55	100.0	239	1	PRIO_AOTHR	P40245	aotus trivi
5	55	100.0	241	1	PRIO_CALMO	P40248	callithrix
6	55	100.0	241	1	PRIO_MANSO	P40255	mandrillus
7	55	100.0	245	1	PRIO_CERAE	P40250	cercopithec
8	55	100.0	246	1	PRIO_CERMO	Q95172	cercopithec
9	55	100.0	246	1	PRIO_CERPA	Q95174	cercopithec
10	55	100.0	246	1	PRIO_CERTO	Q95176	cercopithec
11	55	100.0	252	1	PRIO_ATEPA	P51446	ateles pani
12	55	100.0	252	1	PRIO_CALJA	P40247	callithrix
13	55	100.0	252	1	PRIO_CERAP	P40249	cebun apell
14	55	100.0	252	1	PRIO_RABIT	Q95211	oryctolagus
15	55	100.0	253	1	PRIO_COLGU	P40251	colobus gue
16	55	100.0	253	1	PRIO_GORGO	P40252	gorilla gor
17	55	100.0	253	1	PRIO_HUMAN	P04156	homo sapien
18	55	100.0	253	1	PRIO_MACFA	P40254	macaca fasc
19	55	100.0	253	1	PRIO_PANTR	P40253	pan troglod
20	55	100.0	253	1	PRIO_PONPY	P40256	pongo pygma
21	55	100.0	253	1	PRIO_PREFR	P40257	presbytis f
22	55	100.0	254	1	PRIO_CRIGR	Q60506	cricetus
23	55	100.0	254	1	PRIO_CRIMI	Q60468	cricetus
24	55	100.0	254	1	PRIO_MESAU	P04273	mesocricetu
25	55	100.0	254	1	PRIO_MOUSE	P04925	mus musculus
26	55	100.0	254	1	PRIO_RAT	P13852	rattus norv
27	55	100.0	254	1	PRIO_SIGHI	Q92073	sigmodon hi
28	55	100.0	255	1	PRIO_CANDR	P79141	canamelus dro
29	55	100.0	255	1	PRIO_CANFA	Q46501	canis fami
30	55	100.0	256	1	PRIO_CAPHI	P52113	capra hircu
31	55	100.0	256	1	PRIO_CEREL	P79142	ceruus elap
32	55	100.0	256	1	PRIO_FELCA	O18754	felis silve
33	55	100.0	256	1	PRIO_ODOHE	P47852	odocoileus

RESULT 1

ID	PRIO_ATEGE	STANDARD	PRT	232 AA
AC	P40246			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	01-OCT-1996 (Rel. 34, Last annotation update)			
DE	Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C) (Fragment).			
GN	PRNP.			
OS	Ateles geoffroyi (Black-handed spider monkey).			
OC	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateles.			
OX	NCBI_TaxID=9509;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=95139066; Pubmed=7837269;			
RA	Schatz H.M., DeCosta M., Taylor L., Cohen F.E., Prusiner S.B.;			
RT	"Prion protein gene variation among primates.";			
RL	J. Mol. Biol. 245:362-374(1995).			
CC	- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE			
CC	HOT GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.			
CC	- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED			
CC	"RODS".			
CC	- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.			
CC	- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND			
CC	ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,			
CC	CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME			
CC	(GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),			
CC	TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.			
CC	- SIMILARITY: BELONGS TO THE PRION FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; U08309; AAC50097.1; -			
DR	HSSP; P04156; 1E1G.			
DR	InterPro; IPR000817; Prion.			
DR	Pfam; PF00377; prion; 1.			
DR	SMART; SM00157; PRP; 1.			
DR	PROSITE; PS00291; PRION_1; 1.			
DR	PROSITE; PS00706; PRION_2; 1.			
KW	Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.			
FT	NON_TER 1			
FT	SIGNAL <1	15		BY SIMILARITY.
FT	CHAIN 16	214		MAJOR PRION PROTEIN.
FT	PROPEP 215	>232		REMOVED IN MATURE FORM (BY SIMILARITY).
FT	LIPID 214	214		GPI-ANCHOR (BY SIMILARITY).
FT	DISULFID 163	198		BY SIMILARITY.
FT	CARBOHYD 165	165		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 181	181		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	DOMAIN 44	84		4 X 8 AA TANDEM REPEATS OF P-H-G-G-W-G-

P23907 ovis aries
Q01880 bos taurus
P40243 tragelaphus
P52114 mustela put
P40244 mustela vis
P49927 sus scrofa
P40258 saimiri sci
P10279 bos taurus
P40242 tragelaphus
P51780 trichosurus
P26022 homo sapien
P75995 escherichia

ALIGNMENTS

FT REPEAT 44 51 0.
FT REPEAT 52 59 1.
FT REPEAT 60 67 2.
FT REPEAT 68 75 3.
FT REPEAT 232 232 4.
FT NON_TER 232 232
SQ SEQUENCE 232 AA; 25596 MW; 0E2D75F04C05CC4A CRC64;

Query Match 100.0%; Score 55; DB 1; Length 232;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8
IIIIIIII
Db 48 GWGQPHGG 55

RESULT 2

PRIO_CERAT
ID AC Q95145; Q95200; STANDARD; PRT; 238 AA.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C) (Fragment).
GN PRNP.
OS Cercopithecus aethiops, and
OS Macaca sylvanus (Barbary ape).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=36222, 9546;
RN [1]
RP SEQUENCE FROM N.A.
RA der Kuyt A.C., Dekker J.T., Goudsmit J.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
CC -!- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
CC "RODS".
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
CC CRETZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
CC -!- SIMILARITY: BELONGS TO THE PRION FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; U75384; AAB50623.1; -.
CC EMBL; U75382; AAB50629.1; -.
CC HSSP; P04925; IAG2.
CC InterPro: IPR000817; Prion.
CC Pfam; PF00377; prion; 1.
CC SMART; SM00157; PRP; 1.
CC PROSITE; PS00291; PRION_1; 1.
CC PROSITE; PS00706; PRION_2; 1.
CC Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
FT NON_TER 1 1
FT SIGNAL <1 15 BY SIMILARITY.
FT CHAIN 16 215 MAJOR PRION PROTEIN.
FT PROPEP 216 238 REMOVED IN MATURE FORM (BY SIMILARITY).
FT LIPID 215 215 GPI-ANCHOR (BY SIMILARITY).
FT DISULFID 164 199 BY SIMILARITY.
FT CARBOHYD 166 166 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 182 182 N-LINKED (GLCNAC...) (POTENTIAL).
FT DOMAIN 44 76 4 X 8 AA TANDEM REPEATS OF P-H-G-G-G-W-G-
FT REPEAT 44 52 Q.
FT REPEAT 53 60 1.
FT REPEAT 61 68 2.
FT REPEAT 69 76 3.
FT REPEAT 238 AA; 26123 MW; 5F59A3EBC3E3531B CRC64;
SQ SEQUENCE 238 AA; 26123 MW; 5F59A3EBC3E3531B CRC64;

Query Match 100.0%; Score 55; DB 1; Length 238;
Best Local Similarity 100.0%; Pred. No. 0.035;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8
IIIIIIII
Db 49 GWGQPHGG 56

RESULT 3

PRIO_THEGE
ID AC Q95270; STANDARD; PRT; 238 AA.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C) (Fragment).
GN PRNP OR PRP
OS Theropithecus gelada (Gelada baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Theropithecus.
OX NCBI_TaxID=9565;
RN [1]
RP SEQUENCE FROM N.A.
RA der Kuyt A.C., Dekker J.T., Goudsmit J.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
CC -!- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
CC "RODS".
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
CC CRETZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
CC -!- SIMILARITY: BELONGS TO THE PRION FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; U75383; AAB50630.1; -.
CC HSSP; P04925; IAG2.
CC InterPro: IPR000817; Prion.
CC Pfam; PF00377; prion; 1.
CC SMART; SM00157; PRP; 1.
CC PROSITE; PS00291; PRION_1; 1.
CC PROSITE; PS00706; PRION_2; 1.
CC Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
FT NON_TER 1 1
FT SIGNAL <1 15 BY SIMILARITY.
FT CHAIN 16 >238 MAJOR PRION PROTEIN.
FT DISULFID 164 199 BY SIMILARITY.
FT CARBOHYD 166 166 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 182 182 N-LINKED (GLCNAC...) (POTENTIAL).
FT DOMAIN 44 83 4 X 8 AA TANDEM REPEATS OF P-H-G-G-G-W-G-
FT

FT REPEAT 44 52 1.
 FT REPEAT 53 60 2.
 FT REPEAT 61 68 3.
 FT REPEAT 69 76 4.
 FT NON_TER 238
 SQ SEQUENCE 238 AA; 26104 MW; 5F59BFF602243EDB CRC64;

Query Match 100.0%; Score 55; DB 1; Length 238;
 Best Local Similarity 100.0%; Pred. No. 0.035;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8
 Db 49 GWGQPHGG 56
 |||||

RESULT 4
 ID PRIO_AOTTR STANDARD; PRT; 239 AA.
 AC P40245;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C) (Fragment).
 GN PRNP.
 OS Aotus trivirgatus (Night monkey) (Douroucoulis).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Aotinae; Aotus.
 OX NCBI_TaxID=9505;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95139066; PubMed=7837269;
 RA Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;
 RT "Prion protein gene variation among primates.";
 RL J. Mol. Biol. 245:362-374(1995).
 CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
 CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
 CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
 CC "RODS".
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
 CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
 CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
 CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
 CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
 CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; U08293; AAC50082.1; --
 DR HSSP; P04925; IAG2.
 DR InterPro; IPR000817; Prion.
 DR Pfam; PF00377; prion; 1.
 DR SMART; SM00157; prion; 1.
 DR PROSITE; PS00291; PRION_1; 1.
 DR PROSITE; PS00706; PRION_2; 1.
 KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
 FT NON_TER 1
 FT SIGNAL <1 15
 FT CHAIN 16 >239
 FT MAJOR PRION PROTEIN.
 FT BY SIMILARITY.
 FT FT CHAIN 171 206
 FT BY SIMILARITY.
 FT FT DISULFID 173 173
 FT N-LINKED (GLNAC. . .) (POTENTIAL).
 FT FT CARBOHYD 189 189
 FT N-LINKED (GLNAC. . .) (POTENTIAL).
 FT FT CARBOHYD 44 83
 FT 5 X 8 AA TANDEM REPEATS OF P-H-G-G-G-W-G-
 FT DOMAIN 44
 FT Q.
 FT REPEAT 44 51
 FT Q.

FT REPEAT 52 59 2.
 FT REPEAT 60 67 3.
 FT REPEAT 68 75 4.
 FT REPEAT 76 83 5.
 FT NON_TER 239 239
 SQ SEQUENCE 239 AA; 26246 MW; 2EFB77E354B7024A CRC64;

Query Match 100.0%; Score 55; DB 1; Length 239;
 Best Local Similarity 100.0%; Pred. No. 0.035;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8
 Db 48 GWGQPHGG 55
 |||||

RESULT 5
 ID PRIO_CALMO STANDARD; PRT; 241 AA.
 AC P40248;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C) (Fragment).
 GN PRNP.
 OS Callicebus moloch (Dusky titi).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Callicebinae;
 OC Callicebus.
 OX NCBI_TaxID=9523;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95139066; PubMed=7837269;
 RA Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;
 RT "Prion protein gene variation among primates.";
 RL J. Mol. Biol. 245:362-374(1995).
 CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
 CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
 CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
 CC "RODS".
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
 CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
 CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
 CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
 CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
 CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; U08312; AAC50100.1; --
 DR HSSP; P04925; IAG2.
 DR InterPro; IPR000817; Prion.
 DR Pfam; PF00377; prion; 1.
 DR SMART; SM00157; prion; 1.
 DR PROSITE; PS00291; PRION_1; 1.
 DR PROSITE; PS00706; PRION_2; 1.
 KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
 FT NON_TER 1
 FT SIGNAL <1 15
 FT CHAIN 16 >241
 FT MAJOR PRION PROTEIN.
 FT BY SIMILARITY.
 FT FT CHAIN 172 207
 FT BY SIMILARITY.
 FT FT DISULFID 174 174
 FT N-LINKED (GLNAC. . .) (POTENTIAL).
 FT FT CARBOHYD 190 190
 FT N-LINKED (GLNAC. . .) (POTENTIAL).
 FT FT CARBOHYD 44 84
 FT 5 X 8 AA TANDEM REPEATS OF P-H-G-G-G-W-G-
 FT DOMAIN 44
 FT Q.
 FT REPEAT 44 84
 FT Q.

```
FT REPEAT 44 52 1.
FT REPEAT 53 60 2.
FT REPEAT 61 68 3.
FT REPEAT 69 76 4.
FT REPEAT 77 84 5.
FT NON_TER 241 241
SQ SEQUENCE 241 AA; 26373 MW; C6D2013EE7CAEC93 CRC64;

Query Match 100.0%; Score 55; DB 1; Length 241;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8
Db 57 GWGQPHGG 64

RESULT 6
PRIO_MANSIP
ID PRIO_MANSIP STANDARD; PRT; 241 AA.
AC P40255;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C) (Fragment).
GN PRNP.
OS Mandrilus sphinx (Mandrill) (Papio sphinx).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Cercopithecoidea.
OC NCBI_TaxID=9561;
RN SEQUENCE FROM N.A.
RP MEDLINE=95139066; PubMed=7837269;
RA Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;
RT "Prion protein gene variation among primates.";
RL J. Mol. Biol. 245:362-374(1995).
CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
CC -1- HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
CC "RODS".
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U08303; AAC50091.1; -.
CC HSP; P04925; IAG2.
CC InterPro: IPR000817; Prion.
CC Pfam; PF00377; prion; 1.
CC SMART; SM00157; PRP; 1.
CC PROSITE; PS00291; PRION_1; 1.
CC PROSITE; PS00706; PRION_2; 1.
KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
FT NON_TER 1 15
FT SIGNAL <1 15 BY SIMILARITY.
FT CHAIN 16 223 MAJOR PRION PROTEIN.
FT PROPEP 224 >241 REMOVED IN MATURE FORM (BY SIMILARITY).
FT LIPID 223 223 GPI-ANCHOR (BY SIMILARITY).
FT DISULFD 172 207 BY SIMILARITY.
FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DOMAIN 44 84 5 X 8 AA TANDEN REPEATS OF P-H-G-G-G-W-G-
FT REPEAT 44 52 1.
FT REPEAT 53 60 2.
FT REPEAT 61 68 3.
FT REPEAT 69 76 4.
FT REPEAT 77 84 5.
FT NON_TER 241 241
SQ SEQUENCE 241 AA; 26398 MW; E539D84E2E2B59DE CRC64;

Query Match 100.0%; Score 55; DB 1; Length 241;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8
Db 49 GWGQPHGG 56

RESULT 7
PRIO_CERAE
ID PRIO_CERAE STANDARD; PRT; 245 AA.
AC P40250;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C).
GN PRNP.
OS Cercopithecus aethiops (Green monkey) (Grivet), and
OS Cercopithecus diana (Diana monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Cercopithecoidea.
OC NCBI_TaxID=9534; 36224;
RN SEQUENCE FROM N.A.
RP MEDLINE=95139066; PubMed=7837269;
RA Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;
RT "Prion protein gene variation among primates.";
RL J. Mol. Biol. 245:362-374(1995).
CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
CC -1- HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
CC "RODS".
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U08291; AAC50080.1; -.
CC HSP; U08292; AAC50081.1; -.
CC HSP; P04925; IAG2.
CC InterPro: IPR000817; Prion.
CC Pfam; PF00377; prion; 1.
CC PRINTS; PR00341; PRION.
CC SMART; SM00157; PRP; 1.
CC PROSITE; PS00291; PRION_1; 1.
CC PROSITE; PS00706; PRION_2; 1.
KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
FT SIGNAL 1 22 BY SIMILARITY.
```

FT CHAIN 23 222 MAJOR PRION PROTEIN.
FT PROPEP 223 245 REMOVED IN MATURE FORM. (BY SIMILARITY).
FT LIPID 222 222 GPI-ANCHOR (BY SIMILARITY).
FT DISULFID 171 206 BY SIMILARITY.
FT CARBOHYD 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DOMAIN 51 83 4 X 8 AA TANDEM REPEATS OF P-H-G-G-W-G-
Q.
FT REPEAT 51 59 1.
FT REPEAT 60 67 2.
FT REPEAT 68 75 3.
FT REPEAT 76 83 4.
SQ SEQUENCE 245 AA; 26885 MW; D582B58E2726C99A CRC64;

Query Match 100.0%; Score 55; DB 1; Length 245;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8
|||||||
Db 56 GWGQPHGG 63

RESULT 8
PRIO_CERMO STANDARD; PRT; 246 AA.
AC Q95172; Q95173;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C) (Fragment).
GN PRNP.
OS Cercopithecus mona, and
OS Cercopithecus neglectus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Cercopithecus.
OX NCBI_TaxID=36226, 36227;
RN [1]
RP SEQUENCE FROM N.A.
RA der Kuyil A.C., Dekker J.T., Goudsmit J.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
CC "RODS".
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U75386; AAB50625.1; -;
CC EMBL; U75387; AAB50626.1; -;
CC HSP; P04925; IAG2.
CC InterPro; IPR000817; Prion.
CC Pfam; PF00377; prion; 1.
CC SMART; SM00157; PRP; 1.
CC PROSITE; PS00291; PRION_1; 1.
CC PROSITE; PS00706; PRION_2; 1.
KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
FT NON_TER 1 1

FT SIGNAL <1 15 BY SIMILARITY.
FT CHAIN 16 223 MAJOR PRION PROTEIN.
FT PROPEP 224 246 REMOVED IN MATURE FORM (BY SIMILARITY).
FT LIPID 223 223 GPI-ANCHOR (BY SIMILARITY).
FT DISULFID 172 207 BY SIMILARITY.
FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DOMAIN 44 84 5 X 8 AA TANDEM REPEATS OF P-H-G-G-W-G-
Q.
FT REPEAT 44 52 1.
FT REPEAT 53 60 2.
FT REPEAT 61 68 3.
FT REPEAT 69 76 4.
FT REPEAT 77 84 5.
SQ SEQUENCE 246 AA; 26900 MW; 835D147CA2B4FDD3 CRC64;

Query Match 100.0%; Score 55; DB 1; Length 246;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8
|||||||
Db 49 GWGQPHGG 56

RESULT 9
PRIO_CERPA STANDARD; PRT; 246 AA.
AC Q95174;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C) (Fragment).
GN PRNP.
OS Cercopithecus patas.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Cercopithecus.
OX NCBI_TaxID=27677;
RN [1]
RP SEQUENCE FROM N.A.
RA der Kuyil A.C., Dekker J.T., Goudsmit J.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
CC "RODS".
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U75388; AAB50627.1; -;
CC HSP; P04925; IAG2.
CC InterPro; IPR000817; Prion.
CC Pfam; PF00377; prion; 1.
CC SMART; SM00157; PRP; 1.
CC PROSITE; PS00291; PRION_1; 1.
CC PROSITE; PS00706; PRION_2; 1.
KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
FT NON_TER 1 1

FT SIGNAL <1 15 BY SIMILARITY.
FT CHAIN 16 223 MAJOR PRION PROTEIN.
FT PROPEP 224 246 REMOVED IN MATURE FORM (BY SIMILARITY).
FT LIPID 223 223 GPI-ANCHOR (BY SIMILARITY).
FT DISULFID 172 207 BY SIMILARITY.
FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DOMAIN 44 84 5 X 8 AA TANDEM REPEATS OF P-H-G-G-W-G-
Q.
FT REPEAT 44 52 1.
FT REPEAT 53 60 2.
FT REPEAT 61 68 3.
FT REPEAT 69 76 4.
FT REPEAT 77 84 5.
SQ SEQUENCE 246 AA; 26886 MW; D35D105BEC53108 CRC64;

Query Match 100.0%; Score 55; DB 1; Length 246;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8
IIIIIIII
DB 49 GWGQPHGG 56

RESULT 10
PRIO_CERTO STANDARD; PRT; 246 AA.
AC 095176;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C) (Fragment).
GN PRNP.
OS Cercopithecus torquatus atys (Red-crowned mangabey) (Sooty mangabey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9531;
RN [1]
RP SEQUENCE FROM N.A.
RA der Kuyt A.C., Dekker J.T., Goudsmit J.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
CC -!- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
CC "RODS".
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMAN-STAUSSLER SYNDROME
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
CC -!- SIMILARITY: BELONGS TO THE PRION FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U75385; AAB50628.1; -
DR HSSP; P04925; IAG2.
DR InterPro; IPR000817; Prion.
DR Pfam; PF00377; prion; 1.
DR SMART; SM00157; PRP; 1.
DR PROSITE; PS00291; PRION.1; 1.
DR PROSITE; PS00706; PRION.2; 1.
KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
FT NON_TER 1 1

FT SIGNAL <1 15 BY SIMILARITY.
FT CHAIN 16 223 MAJOR PRION PROTEIN.
FT PROPEP 224 246 REMOVED IN MATURE FORM (BY SIMILARITY).
FT LIPID 223 223 GPI-ANCHOR (BY SIMILARITY).
FT DISULFID 172 207 BY SIMILARITY.
FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DOMAIN 44 84 5 X 8 AA TANDEM REPEATS OF P-H-G-G-W-G-
Q.
FT REPEAT 44 52 1.
FT REPEAT 53 60 2.
FT REPEAT 61 68 3.
FT REPEAT 69 76 4.
FT REPEAT 77 84 5.
SQ SEQUENCE 246 AA; 26914 MW; F58679CBEC5ADC7 CRC64;

Query Match 100.0%; Score 55; DB 1; Length 246;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8
IIIIIIII
DB 49 GWGQPHGG 56

RESULT 11
PRIO_ATEPA STANDARD; PRT; 252 AA.
AC P51446;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C).
GN PRNP.
OS Ateles paniscus (Black spider monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateles.
OX NCBI_TaxID=9510;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95083661; PubMed=7991600;
RA Cervenakova L., Brown P., Goldfarb L.G., Nagle J., Pettrone K.,
RA Rubenstein R., Dubnick M., Gibbs C.J., Gajdusek D.C.;
RT "Infectious amyloid precursor gene sequences in primates used for
RT experimental transmission of human spongiform encephalopathy.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:12159-12162(1994).
CC -!- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
CC -!- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
CC "RODS".
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMAN-STAUSSLER SYNDROME
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
CC -!- SIMILARITY: BELONGS TO THE PRION FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U15164; AAA68634.1; -
DR HSSP; P04156; 1E1G.
DR InterPro; IPR000817; Prion.
DR Pfam; PF00377; prion; 1.
DR PRINTS; PR00341; PRION.


```

DR SMART; SM00157; PRP; 1.
DR PROSITE; PS00291; PRION_1; 1.
DR PROSITE; PS00706; PRION_2; 1.
KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
FT SIGNAL 1 22
FT CHAIN 23 229
FT PROPEP 230 229
FT LIPID 229 229
FT DISULFID 178 213
FT CARBOHYD 180 180
FT CARBOHYD 196 196
FT DOMAIN 51 90
FT REPEAT 51 58
FT REPEAT 59 66
FT REPEAT 67 74
FT REPEAT 75 82
FT REPEAT 83 90
SQ SEQUENCE 252 AA; 27718 MW; 20EA38A42DCC56D1 CRC64;

Query Match 100.0%; Score 55; DB 1; Length 252;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8
| | | | |
DB 55 GWGQPHGG 62

RESULT 12
PRIO_CALJA STANDARD; PRT; 252 AA.
AC P40247;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C).
GN PRNP.
OS Callithrix jacchus (Common marmoset).
OC Eukaryota; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
OC Callithrix.
OC NCBI_TaxID=9483;
RN [1]
RP MEDLINE=95139066; PubMed=7837269;
RA Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;
RT "Prion protein gene variation among primates.";
RL J. Mol. Biol. 245:362-374(1995).
CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
CC "RODS".
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U08304; AAC50092.1; -
CC HSPSP; P04925; IAG2.
CC InterPro; IPR000817; Prion.

```

```

DR Pfam; PF00377; prion; 1.
DR PRINTS; PR00341; PRION.
DR SMART; SM00157; PRP; 1.
DR PROSITE; PS00291; PRION_1; 1.
DR PROSITE; PS00706; PRION_2; 1.
KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
FT SIGNAL 1 22
FT CHAIN 23 229
FT PROPEP 230 229
FT LIPID 229 229
FT DISULFID 178 213
FT CARBOHYD 180 180
FT CARBOHYD 196 196
FT DOMAIN 51 90
FT REPEAT 51 58
FT REPEAT 59 66
FT REPEAT 67 74
FT REPEAT 75 82
FT REPEAT 83 90
SQ SEQUENCE 252 AA; 27639 MW; B2800B60FD5CE664 CRC64;

Query Match 100.0%; Score 55; DB 1; Length 252;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8
| | | | |
DB 55 GWGQPHGG 62

RESULT 13
PRIO_CEBAP STANDARD; PRT; 252 AA.
AC P40249;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C).
GN PRNP.
OS Cebus apella (Brown-capped capuchin).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebus.
OC NCBI_TaxID=9515;
RN [1]
RP MEDLINE=95139066; PubMed=7837269;
RA Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;
RT "Prion protein gene variation among primates.";
RL J. Mol. Biol. 245:362-374(1995).
CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
CC "RODS".
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U08295; AAC50084.1; -
CC HSPSP; P04156; IE1G.

```

DR InterPro; IPR000817; Prion.
 DR Pfam; PF00377; Prion; 1.
 DR PRINTS; PR00341; PRION.
 DR SMART; SM00157; PRP; 1.
 DR PROSITE; PS00291; PRION_1; 1.
 DR PROSITE; PS00706; PRION_2; 1.
 KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
 FT SIGNAL 1 22
 FT CHAIN 23 229 MAJOR PRION PROTEIN.
 FT PROPEP 230 252 REMOVED IN MATURE FORM (BY SIMILARITY).
 FT LIPID 229 229 GPI-ANCHOR (BY SIMILARITY).
 FT DISULFID 178 213 BY SIMILARITY.
 FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DOMAIN 51 90 5 X 8 AA TANDEM REPEATS OF P-H-G-G-W-G-
 FT REPEAT 51 58 Q.
 FT REPEAT 59 66 1.
 FT REPEAT 67 74 2.
 FT REPEAT 75 82 3.
 FT REPEAT 83 90 4.
 FT REPEAT 84 92 5.
 SQ SEQUENCE 252 AA; 27579 MW; A2DFCA0AD26D7821 CRC64;

Query Match 100.0%; Score 55; DB 1; Length 252;
 Best Local Similarity 100.0%; Pred. No. 0.037;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8
 Db 55 GWGQPHGG 62
 |||||

RESULT 14
 PRIO_RABIT
 ID PRIO_RABIT STANDARD; PRT; 252 AA.
 AC Q95211;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C).
 GN PRNP OR PRP.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NEW ZEALAND WHITE;
 RX MEDLINE=97183665; PubMed=9031631;
 RA Loftus B., Rogers M.;
 RT "Characterization of a prion protein (PrP) gene from rabbit; a
 species with apparent resistance to infection by prions.";
 RL Gene 184:215-219(1997).
 CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
 CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
 CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
 CC "RODS".
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
 CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
 CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMAN-STRAUSSLER SYNDROME
 CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
 CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
 CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U28334; AAC48697.1; -
 DR HSP; P04925; IAG2.
 DR InterPro; IPR000817; Prion.
 DR Pfam; PF00377; prion; 1.
 DR PRINTS; PR00341; PRION.
 DR SMART; SM00157; PRP; 1.
 DR PROSITE; PS00291; PRION_1; 1.
 DR PROSITE; PS00706; PRION_2; 1.
 KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
 FT SIGNAL 1 28 POTENTIAL.
 FT CHAIN 29 252 MAJOR PRION PROTEIN.
 FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 178 213 BY SIMILARITY.
 FT DOMAIN 51 92 5 X 8 AA TANDEM REPEATS OF P-H-G-G-W-G-
 FT REPEAT 51 59 Q.
 FT REPEAT 60 67 1.
 FT REPEAT 68 75 2.
 FT REPEAT 76 83 3.
 FT REPEAT 84 92 4.
 FT REPEAT 84 92 5.
 SQ SEQUENCE 252 AA; 27432 MW; 2E177AAF38B23A54 CRC64;

Query Match 100.0%; Score 55; DB 1; Length 252;
 Best Local Similarity 100.0%; Pred. No. 0.037;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8
 Db 56 GWGQPHGG 63
 |||||

RESULT 15
 PRIO_COLGU
 ID PRIO_COLGU STANDARD; PRT; 253 AA.
 AC P40251;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C).
 GN PRNP.
 OS Colobus guereza (Black-and-white colobus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Colobinae;
 OC Colobus.
 OX NCBI_TaxID=33548;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=95139066; PubMed=7837269;
 RA Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;
 RT "Prion protein gene variation among primates.";
 RL J. Mol. Biol. 245:362-374(1995).
 RN [2]
 RP SEQUENCE OF 8-253 FROM N.A.
 RA der Kuyil A.C., Dekker J.T., Goudsmit J.;
 RA Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
 CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
 CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
 CC "RODS".
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
 CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
 CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMAN-STRAUSSLER SYNDROME
 CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
 CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
 CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; U08297; AAC50086.1; -;
DR EMBL; U75389; AAB50624.1; -;
DR HSP; P04925; IAG2.
DR InterPro; IPR00817; Prion.
DR Pfam; PF00377; prion; 1.
DR PRINTS; PR00341; PRION.
DR SMART; SM00157; PRP; 1.
DR PROSITE; PS00291; PRION_1; 1.
DR PROSITE; PS00706; PRION_2; 1.
KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
FT SIGNAL 1 22
FT CHAIN 23 230
FT PROPEP 231 253
FT LIPID 230 230
FT DISULFID 179 214
FT CARBOHYD 181 181
FT CARBOHYD 197 197
FT DOMAIN 51 91
FT REPEAT 51 59
FT REPEAT 60 67
FT REPEAT 68 75
FT REPEAT 76 83
FT REPEAT 84 91
SQ SEQUENCE 253 AA; 27626 MW; 14B17477881F5316 CRC64;
BY SIMILARITY.
MAJOR PRION PROTEIN.
REMOVED IN MATURE FORM (BY SIMILARITY).
GPI-ANCHOR (BY SIMILARITY).
BY SIMILARITY.
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
5 X 8 AA TANDEM REPEATS OF P-H-G-G-G-W-G-Q.

Query Match 100.0%; Score 55; DB 1; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GWGQPHGG 8
Db 56 GWGQPHGG 63
|||||||

Search completed: August 6, 2002, 10:40:50
Job time: 247 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 10:36:43 ; Search time 58.49 Seconds
(without alignments)
23.661 Million cell updates/sec

Title: US-09-543-188A-1
Perfect score: 55
Sequence: 1 GWGQPHGG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 segs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:**

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	55	100.0	97	4	Q16409
2	55	100.0	105	6	Q97697
3	55	100.0	117	11	Q9WUJ1
4	55	100.0	124	6	Q9TU20
5	55	100.0	141	6	Q97905
6	55	100.0	143	6	Q9M217
7	55	100.0	181	6	Q97911
8	55	100.0	185	6	Q97694
9	55	100.0	195	6	Q97903
10	55	100.0	195	6	Q97693
11	55	100.0	200	6	Q97912
12	55	100.0	202	6	Q97908
13	55	100.0	202	6	Q97696
14	55	100.0	204	6	Q97629
15	55	100.0	204	6	Q9TS18
16	55	100.0	204	6	Q9TS17

17	55	100.0	209	6	Q9TV02
18	55	100.0	211	6	Q77787
19	55	100.0	212	6	Q97698
20	55	100.0	213	6	Q9TV04
21	55	100.0	214	6	Q9TV03
22	55	100.0	215	6	Q97904
23	55	100.0	216	6	Q9TV00
24	55	100.0	220	6	Q02825
25	55	100.0	222	6	Q97913
26	55	100.0	223	6	Q97910
27	55	100.0	226	6	Q97907
28	55	100.0	227	6	Q97964
29	55	100.0	227	6	Q97906
30	55	100.0	227	6	Q97909
31	55	100.0	233	4	P78446
32	55	100.0	235	6	Q97695
33	55	100.0	243	11	P97895
34	55	100.0	245	4	Q15216
35	55	100.0	245	6	Q9M207
36	55	100.0	246	4	Q60489
37	55	100.0	253	4	Q9UPL9
38	55	100.0	253	4	Q96E70
39	55	100.0	253	11	Q9Z0T5
40	55	100.0	254	6	Q9TSF8
41	55	100.0	254	11	Q9Z0T4
42	55	100.0	254	11	Q9QYT9
43	55	100.0	256	6	Q02841
44	55	100.0	256	6	Q62670
45	55	100.0	256	6	Q9TV01

ALIGNMENTS

RESULT 1

Q16409
ID Q16409 PRELIMINARY; PRT; 97 AA.
AC Q16409; Q16407;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PRP AMYLOID (PRION PROTEIN) (FRAGMENT).
GN PRNP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92140671; PubMed=1736177;
RA Brown P., Goldfarb L.G., McCombie W.R., Nieto A., Squillacote D.,
RA Sheremata W., Little B.W., Godec M.S., Gibbs C.J. Jr., Gajdusek D.C.;
RT "Atypical Creutzfeldt-Jakob disease in an American family with an
RT insert mutation in the PRNP amyloid precursor gene";
RL Neurology 42:422-427(1992).
RN [2]
RP SEQUENCE OF 1-84 FROM N.A.
RX MEDLINE=96057178; PubMed=7572084;
RA Oda T., Kitamoto T., Tateishi J., Mitsuhashi T., Iwabuchi K., Haga C.,
RA Oguni E., Kato Y., Tomimaga I., Yanai K.;
RT "Prion disease with 144 base pair insertion in a Japanese family
RT line";
RL Acta Neuropathol. 90:80-86(1995).
RN [3]
RP SEQUENCE OF 1-81 FROM N.A.
RX MEDLINE=96390485; PubMed=8797471;
RA Cochran E.J., Bennett D.A., Cervenakova L., Kenney K., Bernard B.,
RA Foster N.L., Benson D.F., Goldfarb L.G., Brown P.;
RT "Familial Creutzfeldt-Jakob disease with a five-repeat octapeptide
RT insert mutation";
RL Neurology 47:727-733(1996).
DR EMBL; S80539; AAB21334.1; -.
DR EMBL; S79978; AAB35416.1; -.

097905
ID O97905 PRELIMINARY; PRT; 141 AA.
AC O97905;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE PRION PROTEIN (FRAGMENT).
GN PRP.
OS Tursiops truncatus (Atlantic bottle-nosed dolphin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Delphinidae;
OC Tursiops.
OX NCBI_TaxID=9739;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PBL;
RX MEDLINE=99303687; PubMed=10373359;
RA Wopfner F., Weidenhofer G., Schneider R., von Brunn A., Gilch S.,
RA Schwarz T.F., Werner T., Schatzl H.M.;
RT "Analysis of 27 mammalian and 9 avian PrPs reveals high conservation
of flexible regions of the prion protein.";
RL J. Mol. Biol. 289:1163-1178(1999).
DR EMBL; AF117311; AAD19982.1; -;
DR HSP; P04925; IAG2.
DR InterPro; IPR000817; Prion.
DR Pfam; PF00377; Prion; 1.
DR PRINTS; PR00341; PRION.
DR SMART; SM00157; PRP; 1.
DR PROSITE; PS00291; PRION_1; 1.
FT NON_TER 1
FT NON_TER 141
SQ SEQUENCE 141 AA; 15066 MW; 79BE306E2AAA187C CRC64;

Query Match 100.0%; Score 55; DB 6; Length 141;
Best Local Similarity 100.0%; Pred. No. 0.097;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8
Db 12 GWGQPHGG 19

RESULT 6
ID Q9MZ17 PRELIMINARY; PRT; 143 AA.
AC Q9MZ17;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE PRION PROTEIN (FRAGMENT).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RA Hermann L.M., Davis W.B., Li H., Wardrop J., Sy M.-S., Gambetti P.,
RA Knowles D.P.;
RT "Detection of PrPc in peripheral blood mononuclear cells of scrapie
susceptible sheep.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF267507; AAF91403.1; -;
DR HSP; P04925; IAG2.
DR InterPro; IPR000817; Prion.
DR PRINTS; PR00341; PRION.
DR SMART; SM00157; PRP; 1.
DR PROSITE; PS00291; PRION_1; 1.
FT NON_TER 143
SQ SEQUENCE 143 AA; 14428 MW; 4E2D296C6C8022E2 CRC64;

Query Match 100.0%; Score 55; DB 6; Length 143;
Best Local Similarity 100.0%; Pred. No. 0.099;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GWGQPHGG 8
Db 59 GWGQPHGG 66
RESULT 7
ID O97911 PRELIMINARY; PRT; 181 AA.
AC O97911;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE PRION PROTEIN (FRAGMENT).
GN PRP.
OS Budorcas taxicolor (takin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Budorcas.
OX NCBI_TaxID=37181;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PBL;
RX MEDLINE=99303687; PubMed=10373359;
RA Wopfner F., Weidenhofer G., Schneider R., von Brunn A., Gilch S.,
RA Schwarz T.F., Werner T., Schatzl H.M.;
RT "Analysis of 27 mammalian and 9 avian PrPs reveals high conservation
of flexible regions of the prion protein.";
RL J. Mol. Biol. 289:1163-1178(1999).
DR EMBL; AF117326; AAD19997.1; -;
DR HSP; P04925; IAG2.
DR InterPro; IPR002395; Kininogen.
DR InterPro; IPR000817; Prion.
DR Pfam; PF00377; Prion; 1.
DR PRINTS; PR00334; KININOGEN.
DR PRINTS; PR00341; PRION.
DR SMART; SM00157; PRP; 1.
DR PROSITE; PS00291; PRION_1; 1.
FT NON_TER 1
FT NON_TER 181
SQ SEQUENCE 181 AA; 19253 MW; A9001D086442E92A CRC64;

Query Match 100.0%; Score 55; DB 6; Length 181;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GWGQPHGG 8
Db 32 GWGQPHGG 39
RESULT 8
ID O97694 PRELIMINARY; PRT; 185 AA.
AC O97694;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE PRION PROTEIN (FRAGMENT).
GN PRP.
OS Cervus nippon dybowskii.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;
OC Cervidae; Cervinae; Cervus.
OX NCBI_TaxID=88065;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97317556; PubMed=9174569;
RA Schatzl H.M., Wopfner F., Gilch S., von Brunn A., Jager G.;

RT "Is codon 129 of prion protein polymorphic in human beings but not in
RL animals?";
RN Lancet 349:1603-1604(1997).

RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-99303687; PubMed-10373359;
RA Wopner F., Weidenhofer G., Schneider R., von Brunn A., Gilch S.,
Schwarz T.F., Werner T., Schatzl H.M.;
RT "Analysis of 27 mammalian and 9 avian PrPs reveals high conservation
of flexible regions of the prion protein.";
RL J. Mol. Biol. 289:1163-1178(1999).
DR EMBL; AF113941; AAD13289.1; -;
DR HSSP; P04925; IAG2.
DR InterPro; IPR002395; Kininogen.
DR InterPro; IPR000817; Prion.
DR Pfam; PF00377; prion; 1.
DR PRINTS; PR00334; KININOGEN.
DR PRINTS; PR00341; PRION.
DR SMART; SM00157; PRP; 1.
DR PROSITE; PS00291; PRION_1; 1.
FT NON_TER 1
FT NON_TER 185
SQ SEQUENCE 185 AA; 19870 MW; BB87C7658BC66E79 CRC64;

Query Match 100.0%; Score 55; DB 6; Length 185;

Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8
Db 35 GWGQPHGG 42

RESULT 9

ID O97903 PRELIMINARY; PRT; 195 AA.
AC O97903;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PRION PROTEIN (FRAGMENT).
GN PRP.
OS Adax nasomaculatus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Hippotraginae; Addax.
OX NCBI_TaxID=59515;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=PBL;
RA MEDLINE-99303687; PubMed-10373359;
RA Wopner F., Weidenhofer G., Schneider R., von Brunn A., Gilch S.,
Schwarz T.F., Werner T., Schatzl H.M.;
RT "Analysis of 27 mammalian and 9 avian PrPs reveals high conservation
of flexible regions of the prion protein.";
RL J. Mol. Biol. 289:1163-1178(1999).
DR EMBL; AF117309; AAD19980.1; -;
DR HSSP; P04925; IAG2.
DR InterPro; IPR002395; Kininogen.
DR InterPro; IPR000817; Prion.
DR Pfam; PF00377; prion; 1.
DR PRINTS; PR00334; KININOGEN.
DR PRINTS; PR00341; PRION.
DR SMART; SM00157; PRP; 1.
DR PROSITE; PS00291; PRION_1; 1.
DR PROSITE; PS00706; PRION_2; 1.
FT NON_TER 1
FT NON_TER 195
SQ SEQUENCE 195 AA; 21321 MW; 6A9BA6A7E1A6CA9 CRC64;

Query Match 100.0%; Score 55; DB 6; Length 195;

Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8
Db 23 GWGQPHGG 30

RESULT 10

ID O97693 PRELIMINARY; PRT; 195 AA.
AC O97693;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PRION PROTEIN (FRAGMENT).
GN PRP.
OS Canis lupus (Gray wolf).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9612;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99303687; PubMed-10373359;
RA Wopner F., Weidenhofer G., Schneider R., von Brunn A., Gilch S.,
Schwarz T.F., Werner T., Schatzl H.M.;
RT "Analysis of 27 mammalian and 9 avian PrPs reveals high conservation
of flexible regions of the prion protein.";
RL J. Mol. Biol. 289:1163-1178(1999).
DR EMBL; AF113939; AAD12063.1; -;
DR HSSP; P04925; IAG2.
DR InterPro; IPR000817; Prion.
DR Pfam; PF00377; prion; 1.
DR PRINTS; PR00341; PRION.
DR SMART; SM00157; PRP; 1.
DR PROSITE; PS00291; PRION_1; 1.
DR PROSITE; PS00706; PRION_2; 1.
FT NON_TER 1
FT NON_TER 195
SQ SEQUENCE 195 AA; 21097 MW; 9D18E4EB9AA5D031 CRC64;

Query Match 100.0%; Score 55; DB 6; Length 195;

Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8
Db 18 GWGQPHGG 25

RESULT 11

ID O97912 PRELIMINARY; PRT; 200 AA.
AC O97912;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PRION PROTEIN (FRAGMENT).
GN PRP.
OS Bison bonasus (European bison).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bison.
OX NCBI_TaxID=9902;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=PBL;
RA MEDLINE-99303687; PubMed-10373359;
RA Wopner F., Weidenhofer G., Schneider R., von Brunn A., Gilch S.,
Schwarz T.F., Werner T., Schatzl H.M.;
RT "Analysis of 27 mammalian and 9 avian PrPs reveals high conservation
of flexible regions of the prion protein.";
RL J. Mol. Biol. 289:1163-1178(1999).

RL J. Mol. Biol. 289:1163-1178(1999).

DR EMBL; AF117328; AAD19999.1; -
DR HSSP; P04156; 1EIG.
DR InterPro; IPR002395; Kininogen.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000817; Prion.
DR Pfam; PF00377; prion; 1.
DR PRINTS; PR00334; KININOGEN.
DR PRINTS; PR00341; PRION.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00157; PRP; 1.
DR PROSITE; PS00291; PRION_1; 1.
DR PROSITE; PS00706; PRION_2; 1.
FT NON_TER 1 1
FT NON_TER 200 200
SQ SEQUENCE 200 AA; 21674 MW; 1F270CDF4BE5271B CRC64;

Query Match 100.0%; Score 55; DB 6; Length 200;

Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8

Db 16 GWGQPHGG 23

RESULT 12

ID O97908 PRELIMINARY; PRT; 202 AA.

AC O97908;

DT 01-MAY-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE PRION PROTEIN (FRAGMENT).

GN PRP.

OS Capra ibex nubiana.

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Caprinae; Capra.

OX NCBI_TaxID=79916;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=PBL;

RX MEDLINE=99303687; PubMed=10373359;

RA Wopner F., Weidenhofer G., Schneider R., von Brunn A., Gilch S.,

RA Schwarz T.F., Werner T., Schatzl H.M.;

RT "Analysis of 27 mammalian and 9 avian PRPs reveals high conservation

of flexible regions of the prion protein.";

RL J. Mol. Biol. 289:1163-1178(1999).

DR EMBL; AF117319; AAD19990.1; -

DR HSSP; P04925; IAG2.

DR InterPro; IPR002395; Kininogen.

DR InterPro; IPR000817; Prion.

DR Pfam; PF00377; prion; 1.

DR PRINTS; PR00334; KININOGEN.

DR PRINTS; PR00341; PRION.

DR SMART; SM00157; PRP; 1.

DR PROSITE; PS00291; PRION_1; 1.

DR PROSITE; PS00706; PRION_2; 1.

FT NON_TER 1 1

FT NON_TER 202 202

SQ SEQUENCE 202 AA; 21949 MW; DB0634A43B4DB77F CRC64;

Query Match

Best Local Similarity 100.0%; Score 55; DB 6; Length 202;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8

Db 31 GWGQPHGG 38

RESULT 13

ID O97696

AC O97696 PRELIMINARY; PRT; 202 AA.

DT 01-MAY-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE PRION PROTEIN (FRAGMENT).

GN PRP.

OS Lama glama (Llama).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.

OX NCBI_TaxID=9844;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99303687; PubMed=10373359;

RA Wopner F., Weidenhofer G., Schneider R., von Brunn A., Gilch S.,

RA Schwarz T.F., Werner T., Schatzl H.M.;

RT "Analysis of 27 mammalian and 9 avian PRPs reveals high conservation

of flexible regions of the prion protein.";

RL J. Mol. Biol. 289:1163-1178(1999).

DR EMBL; AF113943; AAD13291.1; -

DR HSSP; P04925; IAG2.

DR InterPro; IPR000317; Prion.

DR Pfam; PF00377; prion; 1.

DR PRINTS; PR00341; PRION.

DR SMART; SM00157; PRP; 1.

DR PROSITE; PS00291; PRION_1; 1.

DR PROSITE; PS00706; PRION_2; 1.

FT NON_TER 1 1

FT NON_TER 202 202

SQ SEQUENCE 202 AA; 21860 MW; FC45232DB773F354 CRC64;

Query Match

Best Local Similarity 100.0%; Score 55; DB 6; Length 202;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GWGQPHGG 8

Db 11 GWGQPHGG 16

RESULT 14

ID O97629

AC O97629 PRELIMINARY; PRT; 204 AA.

DT 01-MAY-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE PRION PROTEIN (FRAGMENT).

GN PRP.

OS Odocoileus virginianus (white-tailed deer).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;

OC Cervidae; Odocoileinae; Odocoileus.

OX NCBI_TaxID=9874;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=BRAIN;

RA O'Rourke K.I., Miller M.W., Wild M.A., Williams E.S.;

RT "PRP alleles in free ranging and captive white tailed deer (Odocoileus

virginianus).";

RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF091538; AAC69636.1; -

DR HSSP; P04925; IAG2.

DR InterPro; IPR002395; Kininogen.

DR InterPro; IPR000817; Prion.

DR Pfam; PF00377; prion; 1.

DR PRINTS; PR00334; KININOGEN.

DR PRINTS; PR00341; PRION.

DR SMART; SM00157; PRP; 1.

DR PROSITE; PS00291; PRION_1; 1.
DR PROSITE; PS00706; PRION_2; 1.
FT NON_TER 1
FT NON_TER 204
SQ SEQUENCE 204 AA; 22154 MW; CA8AE68F2B49C81E CRC64;

Query Match 100.0%; Score 55; DB 6; Length 204;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8
| | | | |
Db 36 GWGQPHGG 43

RESULT 15

QSTS18 PRELIMINARY; PRT; 204 AA.
AC QSTS18;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PRION PROTEIN (FRAGMENT).
GN PRP.
OS Odocolleus virginianus (white-tailed deer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;
OC Cervidae; Odocolleinae; Odocolleus.
OX NCBI_TaxID=9874;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA O'Rourke K.I., Miller M.W., Wild M.A., Williams F.S.;
RT "PrP alleles in free ranging and captive white tailed deer (Odocolleus virginianus).";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF091559; AAC69627.1; -.
DR HSSP; P04925; IAG2.
DR InterPro; IPR000817; Prion.
DR Pfam; PF00377; prion; 1.
DR PRINTS; PR00341; PRION.
DR SMART; SM00157; PRP; 1.
DR PROSITE; PS00291; PRION_1; 1.
DR PROSITE; PS00706; PRION_2; 1.
FT NON_TER 1
FT NON_TER 204
SQ SEQUENCE 204 AA; 22181 MW; CA962B93FA84D4D3 CRC64;

Query Match 100.0%; Score 55; DB 6; Length 204;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8
| | | | |
Db 36 GWGQPHGG 43

Search completed: August 6, 2002, 10:40:26
Job time: 223 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 10:38:47 ; Search time 73.38 seconds
(without alignments)
9.082 Million cell updates/sec

Title: US-09-543-188a-23

Perfect score: 36

Sequence: 1 IFFWIK 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	100.0	309	22	Mouse T2R22 amino
2	35	97.2	487	22	C glutamicum prote
3	34	94.4	54	22	Human Immune/haema
4	32	88.9	83	22	Human polypeptide
5	32	88.9	99	22	Human digestive sy
6	31	86.1	53	22	Peptide #9009 enco
7	31	86.1	53	22	Protein #7377 enco
8	31	86.1	53	22	Human brain expres
9	31	86.1	53	22	Human bone marrow
10	31	86.1	53	22	Peptide #6962 enco
11	31	86.1	53	22	Peptide #9330 enco

12	31	86.1	79	21	AA13333	Conserved amino ac
13	31	86.1	84	22	AA011977	Human polypeptide
14	31	86.1	114	21	AA060465	Arabidopsis thalia
15	31	86.1	115	22	AA008193	Human polypeptide
16	31	86.1	117	22	AA010930	Human polypeptide
17	31	86.1	118	21	AA060464	Arabidopsis thalia
18	31	86.1	126	22	AA000187	Human polypeptide
19	31	86.1	131	22	AA000417	Human polypeptide
20	31	86.1	136	22	AA032527	Novel human secret
21	31	86.1	139	21	AA05318	Arabidopsis thalia
22	31	86.1	148	21	AA007382	Arabidopsis thalia
23	31	86.1	151	21	AA007381	Arabidopsis thalia
24	31	86.1	164	21	AA007380	Arabidopsis thalia
25	31	86.1	185	22	AA019264	Human G protein-co
26	31	86.1	283	21	AA058345	Arabidopsis thalia
27	31	86.1	287	21	AA058344	Arabidopsis thalia
28	31	86.1	346	20	AA034642	Chlamydia pneumoni
29	31	86.1	476	22	AA023078	Novel human enzyme
30	31	86.1	505	22	AA093525	Human polypeptide
31	31	86.1	582	22	AB069755	Drosophila melano
32	31	86.1	692	21	AA040996	Human ORFX ORF760
33	31	86.1	951	22	AA034872	E. coli cellular p
34	31	86.1	951	22	AA038260	Salmonella typhi c
35	31	86.1	1060	22	AB067403	Drosophila melano
36	31	86.1	1066	22	AB058166	Drosophila melano
37	31	86.1	1229	22	AB05036	Drosophila melano
38	30	83.3	36	21	AA011895	Arabidopsis thalia
39	30	83.3	53	22	AA024024	Arabidopsis EST en
40	30	83.3	62	21	AA044916	Human secreted pro
41	30	83.3	63	22	AA087811	Human immune/haema
42	30	83.3	74	22	AA008609	Human polypeptide
43	30	83.3	106	21	AA034151	Zea mays protein f
44	30	83.3	139	22	AB050596	Novel human diagno
45	30	83.3	139	22	AB063339	Human breast cance

ALIGNMENTS

RESULT 1
AAB87816
ID AAB87816 standard; Protein; 309 AA.
AC AAB87816;
DT 16-MAY-2001 (first entry)
DE Mouse T2R22 amino acid sequence SEQ ID NO:147.

Human; rat; mouse; T2R; taste receptor; G-protein coupled receptor;
taste transduction G-protein coupled receptor; identification; tongue;
taste sensory neuron; taste cell; taste modulator; food;
taste signalling pathway.

OS Mus sp.

PN WO200118050-A2.

XX 15-MAR-2001.

PD 08-SEP-2000; 2000WO-US24821.

XX 10-SEP-1999; 99US-0393634.

XX 22-FEB-2000; 2000US-0510332.

XX (REGC) UNIV CALIFORNIA.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Zuker CS, Adler JE, Ryba N, Mueller K, Hoon M;

XX WPI; 2001-211396/21.

DR N-PSDB; AAF92563.

XX

PT Nucleic acids encoding the T2R family of G-protein coupled taste
 PT receptors, useful for identifying taste modulators that can be used in
 PT food and pharmaceutical industries to customize taste, for e.g. to
 PT decrease the bitter taste of food -

PS Claim 19; Page 234-235; 249pp; English.

XX AAF92502 to AAF92572 represent nucleic acids which encode taste
 CC transduction G-protein coupled receptors designated T2R proteins.
 CC AAB87731 to AAB87824 represent T2R proteins, and AAB87825 to AAB87830
 CC represent T2R family consensus sequences from the present invention.
 CC The T2R proteins are taste modulators. The nucleic acids are useful as
 CC probes for the identification of taste cells, as the nucleic acids are
 CC specifically expressed in taste cells. They also serve as tools for the
 CC generation of taste topographic maps that elucidate the relationship
 CC between the taste cells of the tongue and taste sensory neurons leading
 CC to taste centres in the brain. The taste modulators are useful for
 CC pharmacological and genetic modulation of taste signalling pathways.
 CC Modulatory compounds comprising T2R proteins can therefore be used in
 CC food and pharmaceutical industries to customise taste, for e.g. to
 CC decrease the bitter taste of food or drugs.

SQ Sequence 309 AA;

Query Match 100.0%; Score 36; DB 22; Length 309;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWIK 6
 :|||||
 Db 118 iffwik 123

RESULT 2

AAG92727
 ID AAG92727 standard; Protein; 487 AA.

XX AAG92727;

XX 26-SEP-2001 (first entry)

XX C glutamicum protein fragment SEQ ID NO: 6481.

XX Corynebacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis.

XX Corynebacterium glutamicum.

XX EP1108790-A2.

XX 20-JUN-2001.

XX 18-DEC-2000; 2000EP-0127688.

XX 16-DEC-1999; 99JP-0377484.

XX 07-APR-2000; 2000JP-0159162.

XX 03-AUG-2000; 2000JP-0280988.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX WPI; 2001-376931/40.

XX N-PSDB; AAH67946.

XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analysing
 PT expression profile or pattern of a gene and identifying homologous gene

XX Claim 17; SEQ ID NO: 6481; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein
 CC sequences from the Corynebacterium glutamicum glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of corynebacterium, measuring expression amount and
 CC analysing the expression profile or expression pattern of a gene derived
 CC from corynebacterium, and identifying a homologue of a gene derived
 CC from corynebacterium. Corynebacterium bacteria are useful for producing
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described
 CC in the exemplification of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.

SQ Sequence 487 AA;

Query Match 97.2%; Score 35; DB 22; Length 487;
 Best Local Similarity 83.3%; Pred. No. 2.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWIK 6
 :|||||
 Db 412 vffwik 417

RESULT 3

AAM83561
 ID AAM83561 standard; Protein; 54 AA.

XX AAM83561;

XX 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen SEQ ID NO:11154.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis.

XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US01354.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216647.

XX 07-JUL-2000; 2000US-0216880.

XX 11-JUL-2000; 2000US-0217487.

XX 14-JUL-2000; 2000US-0218290.

XX 26-JUL-2000; 2000US-0220964.

XX 14-AUG-2000; 2000US-0224518.

XX 14-AUG-2000; 2000US-0224519.

XX 14-AUG-2000; 2000US-0225213.

XX 14-AUG-2000; 2000US-0225214.

XX 14-AUG-2000; 2000US-0225266.

XX 14-AUG-2000; 2000US-0225267.

XX 14-AUG-2000; 2000US-0225268.

Query Match 94.4%; Score 34; DB 22; Length 54;
 Best Local Similarity 83.3%; Pred. No. 41;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IFFWIK 6
 Db 27 Iffwik 32

RESULT 4
 ID AAO05208
 XX AAO05208 standard; Protein; 83 AA.

AC AAO05208;

DT 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 19100.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.

OS Homo sapiens.

XX WO200164835-A2.

PN 07-SEP-2001.

PD 26-FEB-2001; 2001WO-US04927.

PF 28-FEB-2000; 2000US-0515126.

XX 18-MAY-2000; 2000US-0577409.

PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

PI WPI; 2001-514838/56.

XX N-PSDB; AAI85139.

PT Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -

PS Claim 20; SEQ ID NO 19100; 1399pp + Sequence Listing; English.

CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
 CC the encoded proteins (AAO0010-AAO1310) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.

CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 83 AA;

Query Match 88.9%; Score 32; DB 22; Length 83;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 FFWIK 6
 Db 20 ffwik 24

RESULT 5

AAM92732

ID AAM92732 standard; Protein; 99 AA.

XX AAM92732;

DT 06-NOV-2001 (first entry)

DE Human digestive system antigen SEQ ID NO: 2081.

XX Human; digestive system antigen; gene therapy; cancer; appendicitis;
 KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
 KW digestive system disorder; Meckel's diverticulum.

OS Homo sapiens.

XX WO200155314-A2.

PN 02-AUG-2001.

PD 17-JAN-2001; 2001WO-US01324.

PF 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216647.

XX 07-JUL-2000; 2000US-0216880.

XX 11-JUL-2000; 2000US-0217487.

XX 11-JUL-2000; 2000US-0217496.

XX 14-JUL-2000; 2000US-0218290.

XX 26-JUL-2000; 2000US-0220963.

XX 26-JUL-2000; 2000US-0220964.

XX 14-AUG-2000; 2000US-0224518.

XX 14-AUG-2000; 2000US-0224519.

XX 14-AUG-2000; 2000US-0225213.

XX 14-AUG-2000; 2000US-0225214.

XX 14-AUG-2000; 2000US-0225266.

XX 14-AUG-2000; 2000US-0225267.

XX 14-AUG-2000; 2000US-0225268.

XX 14-AUG-2000; 2000US-0225270.

XX 14-AUG-2000; 2000US-0225447.

XX 14-AUG-2000; 2000US-0225757.

XX 14-AUG-2000; 2000US-0225758.

XX 18-AUG-2000; 2000US-0225759.

XX 18-AUG-2000; 2000US-0226279.

XX 22-AUG-2000; 2000US-0226681.

XX 22-AUG-2000; 2000US-0226688.

XX 22-AUG-2000; 2000US-0227182.

XX 23-AUG-2000; 2000US-0227009.

XX 30-AUG-2000; 2000US-0228924.

XX 01-SEP-2000; 2000US-0229287.

XX 01-SEP-2000; 2000US-0229343.

XX 01-SEP-2000; 2000US-0229344.

XX 01-SEP-2000; 2000US-0229345.

XX 05-SEP-2000; 2000US-0229509.

XX 05-SEP-2000; 2000US-0229513.

XX 06-SEP-2000; 2000US-0230437.

XX 06-SEP-2000; 2000US-0230438.

XX 08-SEP-2000; 2000US-0231242.

XX 08-SEP-2000; 2000US-0231243.

XX 08-SEP-2000; 2000US-0231244.

XX 08-SEP-2000; 2000US-0231413.

XX 08-SEP-2000; 2000US-0231414.

PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
PT
XX Claim 27; SEQ ID NO 34138; 639pp + sequence listing; English.
XX The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC foetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 53 AA;

Query Match 86.1%; Score 31; DB 22; Length 53;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IFFWI 5
Db 38 iffwi 42

RESULT 7
ABB25378
ID ABB25378 standard; Protein; 53 AA.
XX
AC ABB25378;
XX
DT 23-JAN-2002 (first entry)
XX Protein #7377 encoded by probe for measuring heart cell gene expression.
XX
DE Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00666.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human

PT hearts -
XX Claim 15; SEQ ID NO 27148; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 53 AA;

Query Match 86.1%; Score 31; DB 22; Length 53;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IFFWI 5
Db 38 iffwi 42

RESULT 8
AAM62374
ID AAM62374 standard; Protein; 53 AA.
XX
AC AAM62374;
XX
DT 05-NOV-2001 (first entry)
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 34479.
XX
DE Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 34479; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,

CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.

XX SQ Sequence 53 AA;

Query Match 86.1%; Score 31; DB 22; Length 53;

Best Local Similarity 100.0%; Pred. No. 1.3e+02; Length 53;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IFFWI 5

|||||

Db 38 iffwi 42

RESULT 9

AAM75181

ID AAM75181 standard; Protein; 53 AA.

XX AC AAM75181;

XX DT 06-NOV-2001 (first entry)

XX DE Human bone marrow expressed probe encoded protein SEQ ID NO: 35487.

XX KW Human; bone marrow expressed exon; gene expression analysis; probe;

XX KW microarray; cancer; leukaemia; lymphoma; myeloma.

XX OS Homo sapiens.

XX PN WO200157276-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000668.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488900/53.

XX PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -

XX PS Example 4; SEQ ID NO: 35487; 658pp + Sequence Listing; English.

XX CC The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukaemia and myeloma. The present sequence is a
XX protein encoded by one of the probes of the invention.

XX SQ Sequence 53 AA;

Query Match 86.1%; Score 31; DB 22; Length 53;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IFFWI 5

|||||

Db 38 iffwi 42

RESULT 10

AAM20528

ID AAM20528 standard; Protein; 53 AA.

XX AC AAM20528;

XX DT 12-OCT-2001 (first entry)

XX DE Peptide #6962 encoded by probe for measuring cervical gene expression.

XX KW Probe; human; microarray; gene expression; cervical epithelial cell;

XX KW cervical cancer.

XX OS Homo sapiens.

XX PN WO200157278-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000670.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488901/53.

XX PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human cervical epithelial cells -

XX PS Claim 27; SEQ ID No 25354; 487pp; English.

XX CC The present invention relates to human single exon nucleic acid probes
XX (SENPs; see AAI10068-AAI28459). The present sequence is a peptide encoded
XX by one such probe. The SENPs are derived from human HeLa cells. The SENPs
XX can be used to produce a single exon microarray, which can be used for
XX measuring human gene expression in a sample derived from human cervical
XX epithelial cells. By measuring gene expression, the probes are therefore
XX useful in grading and/or staging of diseases of the cervix, notably
XX cervical cancer.

XX CC Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 53 AA;

Query Match 86.1%; Score 31; DB 22; Length 53;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IFFWI 5

|||||

Db 38 iffwi 42

RESULT 11

AAM35293

ID AAM35293 standard; Protein; 53 AA.

XX AC AAM35293;

XX

```
DT 17-OCT-2001 (first entry)
XX Peptide #9330 encoded by probe for measuring placental gene expression.
DE Probe: microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
KW Homo sapiens.
OS WO200157272-A2.
PN 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US00663.
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-488897/53.
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX Claim 27; SEQ ID No 35562; 654pp; English.
XX The present invention relates to single exon nucleic acid probes (SENP:
CC see AA131315-AA157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.
XX Sequence 53 AA;
SQ
Query Match 86.1%; Score 31; DB 22; Length 53;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IFFWI 5
DB 38 IFFWI 42
RESULT 12
AAB13333
ID AAB13333 standard; Protein; 79 AA.
XX AAB13333;
XX 12-JAN-2001 (first entry)
XX Conserved amino acid sequence #1.
XX Caenorhabditis elegans; human; daf-18; insulin signalling pathway;
KW daf-2; age-1; insulin receptor; PI 3-kinase; PKB kinase;
KW PTEN lipid phosphatase; antidiabetic; anorectic; obesity; diabetes.
XX Caenorhabditis elegans.
OS Homo sapiens.
XX WO200033068-A1.
PN 08-JUN-2000.
PD
```

```
XX 02-DEC-1999; 99WO-US28529.
XX 03-DEC-1998; 98US-0205658.
XX (GEHO ) GEN HOSPITAL CORP.
XX Ruvkun G, Ogg S;
XX WPI; 2000-423022/36.
XX Diagnosing and treating obesity and impaired glucose tolerance using
PT modulators of daf-18 expression and/or activity -
XX Disclosure; Page 398; 402pp; English.
XX The present sequence is a conserved sequence from Caenorhabditis elegans
CC and Homo sapiens. A number of C. elegans proteins have been identified
CC that have mammalian homologues acting in the insulin signalling pathway.
CC The C. elegans age-1 gene encodes a homologue of the mammalian PI
CC 3-kinase whilst daf-2 encodes a homologue of the mammalian insulin
CC receptor. The C. elegans AKT kinase and PKB kinase act downstream of
CC daf-2 and age-1, just as their mammalian homologues act downstream of
CC insulin signalling. The C. elegans PTEN lipid phosphatase homologue,
CC DAF-18, has been found to act upstream of AKT in the pathway. This
CC discovery has enabled mammalian PTEN action to be mapped to the insulin
CC signalling pathway. Conserved DAF motifs can be used to design probes to
CC identify mammalian DAF homologues and thus to identify individuals with a
CC predisposition towards the development of glucose intolerance conditions,
CC such as obesity and diabetes.
XX Sequence 79 AA;
SQ
Query Match 86.1%; Score 31; DB 21; Length 79;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 IFFWIK 6
DB 61 IFFWIK 66
RESULT 13
AAO11977
ID AAO11977 standard; Protein; 84 AA.
XX AAO11977;
XX 06-NOV-2001 (first entry)
XX Human polypeptide SEQ ID NO 25869.
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX Homo sapiens.
XX WO200164835-A2.
XX 07-SEP-2001.
XX 26-FEB-2001; 2001WO-US04927.
XX 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT;
XX
```

DR WPI; 2001-514838/56.
XX N-PSDB; AAI91908.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
XX
PS Claim 20; SEQ ID NO 25869; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI93841) and
CC the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 84 AA;

Query Match 86.1%; Score 31; DB 22; Length 84;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FFWIK 6
|||:
Db 12 ffwvk 16

RESULT 14
AAG60465
ID AAG60465 standard; Protein; 114 AA.
XX
AC AAG60465;
XX
XX 18-OCT-2000 (first entry)
DT
DE Arabidopsis thaliana protein fragment SEQ ID NO: 78318.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
PN
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 21-MAY-1999; 99US-0135629.
PR 24-MAY-1999; 99US-0136021.
PR 25-MAY-1999; 99US-0136392.
PR 27-MAY-1999; 99US-0136782.
PR 28-MAY-1999; 99US-0137222.
PR 01-JUN-1999; 99US-0137528.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.

CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 115 AA;

Query Match 86.1%; Score 31; DB 22; Length 115;
Best Local Similarity 80.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FFWIK 6
|||:
Db 86 ffwvk 90

Search completed: August 6, 2002, 10:38:48
Job time: 125 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 10:39:20 ; Search time 26.75 Seconds
(without alignments)
5.479 Million cell updates/sec

Title: US-09-543-188A-23

Perfect score: 36

Sequence: 1 IFFWIK 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	94.4	3287	2	US-08-477-451-7
2	30	83.3	218	2	US-08-336-031-4
3	30	83.3	218	5	PCT-US95-06725-4
4	30	83.3	411	2	US-08-336-031-2
5	30	83.3	411	2	US-08-902-853-7
6	30	83.3	411	5	PCT-US95-06725-2
7	30	83.3	423	1	US-08-844-064-7
8	30	83.3	423	3	US-09-009-433-7
9	30	83.3	470	4	US-08-068-392-2
10	30	83.3	470	4	US-08-396-988-2
11	30	83.3	553	1	US-08-475-894-2
12	30	83.3	553	1	US-08-484-710-2
13	30	83.3	553	2	US-08-484-709-2
14	30	83.3	553	4	US-08-474-697-2
15	30	83.3	827	4	US-08-669-286-11
16	30	83.3	827	4	US-09-469-253-11
17	30	83.3	827	4	US-09-642-146-11
18	30	83.3	876	1	US-08-785-429-2
19	30	83.3	876	3	US-08-956-621-2
20	30	83.3	883	2	US-08-953-492-2
21	29	80.6	142	2	US-08-678-194-8
22	29	80.6	142	4	US-08-890-011-8
23	29	80.6	142	4	US-09-262-724-8
24	29	80.6	270	4	US-09-085-305-19
25	29	80.6	505	3	US-08-747-221B-14
26	29	80.6	505	4	US-09-005-051-14
27	29	80.6	530	3	US-08-747-221B-53

28 29 80.6 530 4 US-09-005-051-53 Sequence 53, Appl
29 29 80.6 550 3 US-08-747-221B-19 Sequence 19, Appl
30 29 80.6 550 3 US-08-747-221B-58 Sequence 58, Appl
31 29 80.6 550 4 US-09-005-051-19 Sequence 19, Appl
32 29 80.6 550 4 US-09-005-051-58 Sequence 58, Appl
33 29 80.6 1298 1 US-08-222-616-33 Sequence 33, Appl
34 29 80.6 1298 1 US-08-340-011-2 Sequence 2, Appl
35 29 80.6 1298 3 US-08-901-710-2 Sequence 2, Appl
36 29 80.6 1298 5 PCT-US95-04228-33 Sequence 33, Appl
37 29 80.6 1362 2 US-08-874-678-33 Sequence 33, Appl
38 29 80.6 1362 3 US-08-643-839-33 Sequence 33, Appl
39 29 80.6 1363 1 US-08-340-011-4 Sequence 4, Appl
40 29 80.6 1363 2 US-08-874-678-32 Sequence 32, Appl
41 29 80.6 1363 3 US-08-643-839-32 Sequence 32, Appl
42 29 80.6 1363 3 US-08-901-710-4 Sequence 4, Appl
43 29 80.6 1368 2 US-08-874-678-34 Sequence 34, Appl
44 29 80.6 1368 3 US-08-643-839-34 Sequence 34, Appl
45 28 77.8 231 4 US-08-448-489-19 Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-08-477-451-7
: Sequence 7, Application US/08477451
: Patent No. 5928865
: GENERAL INFORMATION:
: APPLICANT: Covacci, Antonello
: TITLE OF INVENTION: Helicobacter Pylori Cagl Region
: NUMBER OF SEQUENCES: 46
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Chiron Corporation
: STREET: 4560 Horton Street
: CITY: Emeryville
: STATE: CA
: COUNTRY: USA
: ZIP: 94608-2916
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/477,451
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: McClung, Barbara G.
: REGISTRATION NUMBER: 33,113
: REFERENCE/DOCKET NUMBER: 0335.002
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 510-601-2708
: TELEFAX: 510-655-3542
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3287 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-477-451-7

Query Match 94.4%; Score 34; DB 2; Length 3287;
Best Local Similarity 83.3%; Pred. No. 4.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IFFWIK 6
| | | | |
Db 679 IFFWIK 684

RESULT 2
US-08-336-031-4
; Sequence 4, Application US/08336031
; Patent No. 5817782
; GENERAL INFORMATION:
; APPLICANT: Jazwinski, S. M.
; TITLE OF INVENTION: LAG1: A GENE FOR INCREASING THE
; FILING DATE: 03-JUN-1994
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336.031
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/253,875
; FILING DATE: 03-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 93032
; TELEPHONE: (516) 742-4366
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 218 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-06725-4

Query Match 83.3%; Score 30; DB 2; Length 218;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IFFWI 5
Db 118 VFFWI 122

RESULT 3
PCT-US95-06725-4
; Sequence 4, Application PC/TUS9506725
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: LAG1: A GENE FOR INCREASING THE
; FILING DATE: 03-JUN-1994
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/06725
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/253,875 & 08/336.031
; FILING DATE: 03-JUN-1994 & 08-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 93032
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4366
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 218 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-06725-4

Query Match 83.3%; Score 30; DB 5; Length 218;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IFFWI 5
Db 118 VFFWI 122

RESULT 4
US-08-336-031-2
; Sequence 2, Application US/08336031
; Patent No. 5817782
; GENERAL INFORMATION:
; APPLICANT: Jazwinski, S. M.
; TITLE OF INVENTION: LAG1: A GENE FOR INCREASING THE
; FILING DATE: 03-JUN-1994
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336.031
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/253,875
; FILING DATE: 03-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 93032
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4366
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids

;
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-336-031-2

Query Match 83.3%; Score 30; DB 2; Length 411;
Best Local Similarity 80.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IFFWI 5
Db 311 VFFWI 315

RESULT 5
US-08-902-853-7
; Sequence 7, Application US/08902853
; Patent No. 5945330
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: HUMAN LONGEVITY-ASSURANCE PROTEIN HOMOLOGS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/902,853
; FILING DATE: Herewith
; CLASSIFICATION: ?
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0345 US
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 541568
US-08-902-853-7

Query Match 83.3%; Score 30; DB 2; Length 411;
Best Local Similarity 80.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IFFWI 5
Db 311 VFFWI 315

RESULT 6
PCT-US95-06725-2
; Sequence 2, Application PC/TUS9506725
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: LAG1: A GENE FOR INCREASING THE
; TITLE OF INVENTION: LONGEVITY OF EUKARYOTES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/06725
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/253,875 & 08/336,031
; FILING DATE: 03-JUN-1994 & 08-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9303Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-06725-2

Query Match 83.3%; Score 30; DB 5; Length 411;
Best Local Similarity 80.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IFFWI 5
Db 311 VFFWI 315

RESULT 7
US-08-844-064-7
; Sequence 7, Application US/08844064
; Patent No. 5747314
; GENERAL INFORMATION:
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: NO. 5747314el Compounds
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS

;; SOFTWARE: FastSEQ for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/844,064
;; FILING DATE: 18-APR-1997
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 9607991.8
;; FILING DATE: 18-APR-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Gimmi, Edward R
;; REGISTRATION NUMBER: 38,891
;; REFERENCE/DOCKET NUMBER: P31458-4
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 610-270-4478
;; TELEFAX: 610-270-5090
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 423 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-844-064-7

Query Match 83.3%; Score 30; DB 1; Length 423;
Best Local Similarity 80.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFW 5

Db 279 IFFW 283

RESULT 8
US-09-009-433-7
; Sequence 7, Application US/09009433
; Patent No. 6087142
; GENERAL INFORMATION:
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: No. 6087142el Compounds
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/844,064
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: 9607991.8
; FILING DATE: 18-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31458-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:

;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 423 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-09-009-433-7

Query Match 83.3%; Score 30; DB 3; Length 423;
Best Local Similarity 80.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFW 5

Db 279 IFFW 283

RESULT 9
US-08-068-392-2
; Sequence 2, Application US/08068392
; Patent No. 6150152
; GENERAL INFORMATION:
; APPLICANT: Shapiro, Steven M.
; TITLE OF INVENTION: Human Macrophage Metalloproteinase
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scott J. Meyer, Monsanto Co., A3SM
; STREET: 800 N. Lindbergh Blvd.
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63167
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/068,392
; FILING DATE: 19930528
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyer, Scott J.
; REGISTRATION NUMBER: 25275
; REFERENCE/DOCKET NUMBER: 07-24(12406)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)694-3117
; TELEFAX: (314)694-5435
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 470 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-068-392-2

Query Match 83.3%; Score 30; DB 4; Length 470;
Best Local Similarity 80.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 FFWIK 6

Db 305 FFWLK 309

RESULT 10
US-08-396-988-2
; Sequence 2, Application US/08396988
; Patent No. 6204043
; GENERAL INFORMATION:
; APPLICANT: Shapiro, Steven M.

;; TITLE OF INVENTION: Human Macrophage Metalloproteinase
;; NUMBER OF SEQUENCES: 3
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Scott J. Meyer, Monsanto Co., A3SM
;; STREET: 800 N. Lindbergh Blvd.
;; CITY: St. Louis
;; STATE: Missouri
;; COUNTRY: USA
;; ZIP: 63167
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/396,988
;; FILING DATE: 01-MAR-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/068,392
;; FILING DATE: 28-MAY-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Meyer, Scott J.
;; REGISTRATION NUMBER: 25275
;; REFERENCE/DOCKET NUMBER: 07-24(12406)A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (314)694-3117
;; TELEFAX: (314)694-5435
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 470 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-396-988-2

Query Match 83.3%; Score 30; DB 4; Length 470;
Best Local Similarity 80.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0;

OY 2 FFWIK 6
DB 305 FFWLK 309

RESULT 11
US-08-475-894-2
;; Sequence 2, Application US/08475894
;; Patent No. 5641748
;; GENERAL INFORMATION:
;; APPLICANT: Yen-Ming Hsu
;; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
;; NUMBER OF SEQUENCES: 6
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: LAHIVE & COCKFIELD
;; STREET: 60 State Street, Suite 510
;; CITY: Boston
;; STATE: Massachusetts
;; COUNTRY: USA
;; ZIP: 02109-1875
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/475,894
;; FILING DATE:
;; CLASSIFICATION: 530
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Louis Myers
;; REGISTRATION NUMBER: 35,965

;; REFERENCE/DOCKET NUMBER: BGP-191
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (617)227-7400
;; TELEFAX: (617)227-5941
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 553 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; FRAGMENT TYPE: internal
US-08-475-894-2

Query Match 83.3%; Score 30; DB 1; Length 553;
Best Local Similarity 80.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0;

OY 1 IFFWI 5
DB 68 IFFWV 72

RESULT 12
US-08-484-710-2
;; Sequence 2, Application US/08484710
;; Patent No. 5656438
;; GENERAL INFORMATION:
;; APPLICANT: Yen-Ming Hsu
;; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
;; NUMBER OF SEQUENCES: 6
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: LAHIVE & COCKFIELD
;; STREET: 60 State Street, Suite 510
;; CITY: Boston
;; STATE: Massachusetts
;; COUNTRY: USA
;; ZIP: 02109-1875
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/484,710
;; FILING DATE:
;; CLASSIFICATION: 530
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Louis Myers
;; REGISTRATION NUMBER: 35,965
;; REFERENCE/DOCKET NUMBER: BGP-190
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (617)227-7400
;; TELEFAX: (617)227-5941
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 553 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; FRAGMENT TYPE: internal
US-08-484-710-2

Query Match 83.3%; Score 30; DB 1; Length 553;
Best Local Similarity 80.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0;

OY 1 IFFWI 5
DB 68 IFFWV 72

RESULT 13
US-08-484-709-2
; Sequence 2, Application US/08484709
; Patent No. 5837844
; GENERAL INFORMATION:
; APPLICANT: Yen-Ming Hsu
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,709
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Louis Myers
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-192
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 553 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-484-709-2

Query Match 83.3%; Score 30; DB 2; Length 553;
Best Local Similarity 80.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IFFWI 5
Db 68 IFFWV 72

RESULT 14
US-08-474-697-2
; Sequence 2, Application US/08474697
; Patent No. 6171800
; GENERAL INFORMATION:
; APPLICANT: Yen-Ming Hsu
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,697
; FILING DATE:

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Louis Myers
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 553 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-474-697-2

Query Match 83.3%; Score 30; DB 4; Length 553;
Best Local Similarity 80.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IFFWI 5
Db 68 IFFWV 72

RESULT 15
US-08-669-286-11
; Sequence 11, Application US/08669286
; Patent No. 6130060
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, SEIJI
; APPLICANT: SAKURAI, TAKASHI
; APPLICANT: NEZU, JUNI-ICHI
; TITLE OF INVENTION: GENE ENCODING ADSEVERIN
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/669,286
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 230-110P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 827 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-669-286-11

Query Match 83.3%; Score 30; DB 4; Length 827;
Best Local Similarity 80.0%; Pred. No. 6.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IFFWI 5
:||||
Db 658 VFFWI 662

Search completed: August 6, 2002, 10:39:22
Job time: 159 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 10:37:27 ; Search time 35:27 seconds
(without alignments)
16.346 Million cell updates/sec

Title: US-09-543-188a-23

Perfect score: 36

Sequence: 1 IFFWIK 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR-71.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	91.7	101	2 B87530	hypothetical prote
2	32	88.9	35	2 D81622	hypothetical prote
3	32	88.9	73	2 S59082	hypothetical prote
4	32	88.9	374	2 T32207	hypothetical prote
5	32	88.9	423	2 E90569	conserved hypothet
6	32	88.9	544	2 T51162	aminoglycoside ace
7	32	88.9	660	2 AE3226	conserved hypothet
8	32	88.9	988	2 H71338	conserved hypothet
9	31	86.1	86	2 AD0850	doubtful CDS found
10	31	86.1	162	2 D69381	hypothetical prote
11	31	86.1	247	2 T33469	hypothetical prote
12	31	86.1	258	2 AG2542	hypothetical prote
13	31	86.1	262	2 S75891	hypothetical prote
14	31	86.1	267	2 A72338	hemolysin - Thermo
15	31	86.1	287	2 T05338	hypothetical prote
16	31	86.1	388	2 S57526	cellulase - Fibrob
17	31	86.1	398	2 H70142	hypothetical prote
18	31	86.1	400	2 H97550	general l-amino ac
19	31	86.1	400	2 AH2770	hypothetical prote
20	31	86.1	449	2 C86496	hypothetical prote
21	31	86.1	449	2 D72127	hypothetical prote
22	31	86.1	449	2 A81544	hypothetical prote
23	31	86.1	474	2 S41117	hypothetical prote
24	31	86.1	495	1 S25942	nitric-oxide reduc
25	31	86.1	497	1 S53834	NADH dehydrogenase
26	31	86.1	500	2 F88921	NADH dehydrogenase
27	31	86.1	507	1 A30828	protein F56E10.3 [
28	31	86.1	513	2 B70478	steroid 17alpha-mo
29	31	86.1	532	2 G82872	probable sodium sy
					cardiolipin synthase

30	31	86.1	564	2 T20191	hypothetical prote
31	31	86.1	597	2 T27497	hypothetical prote
32	31	86.1	639	1 S45776	uracil transport p
33	31	86.1	669	2 I54205	galactosylceramida
34	31	86.1	738	2 S10859	membrane protein t
35	31	86.1	826	1 A31822	villin - chicken
36	31	86.1	951	1 SYECVT	valine-tRNA ligas
37	31	86.1	951	2 C91283	valine tRNA synthet
38	31	86.1	951	2 E86124	valine-tRNA ligas
39	31	86.1	951	2 AC1061	valine-tRNA ligas
40	31	86.1	955	2 F84972	valine-tRNA ligas
41	31	86.1	963	2 T12087	H+-transporting AT
42	31	86.1	965	2 AE0418	valine-tRNA ligas
43	31	86.1	966	2 T52413	H+-transporting AT
44	31	86.1	1101	2 T33153	hypothetical prote
45	31	86.1	1113	2 S48495	probable membrane

ALIGNMENTS

RESULT 1

B87530

hypothetical protein CC2267 [imported] - Caulobacter crescentus

C;Species: Caulobacter crescentus

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C;Accession: B87530

R;Niernan, W.C.; Feldhlyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of Caulobacter crescentus.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: B87530

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-101 <STO>

A;Cross-references: GB:AE005673; NID:g13423780; PIDN:AAK24238.1; GSPDB:GN00148

C;Genetics:

A;Gene: CC2267

Query Match 91.7%; Score 33; DB 2; Length 101;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWIK 6

DB 6 VFFWLK 11

RESULT 2

D81622

hypothetical protein CP0010 [imported] - Chlamydomophila pneumoniae (strain AR39)

C;Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae

C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000

C;Accession: D81622

R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke, C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000

A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39
A;Reference number: A81500; MUID:20150255

A;Accession: D81622

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-35 <REA>

A;Cross-references: GB:AE002165; GB:AE002161; NID:g7188948; PIDN:AAF37906.1; PID:g718

A;Experimental source: strain AR39, HL cells

C;Genetics:

A;Gene: CP0010

Query Match 88.9%; Score 32; DB 2; Length 35;

Best Local Similarity 83.3%; Pred. No. 13;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IFFWIK 6
Db 12 IFFWIE 17

RESULT 3
S59082
hypothetical protein 73 - red alga (Chondrus crispus) mitochondrion
C:Species: mitochondrion Chondrus crispus (carragheen)
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 21-Jul-2000
C:Accession: S59082
R:Leblanc, C.; Boyen, C.; Richard, O.; Bonnard, G.; Grienenberger, J.M.; Kloareg, B.
J. Mol. Biol. 250, 484-495, 1995
A:Title: Complete sequence of the mitochondrial DNA of the rhodophyte Chondrus crispus
A:Reference number: S59078; MUID:95341681
A:Accession: S59082
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-73 <LEB>
A:Cross-references: EMBL:247547; NID:gl019057; PIDN:CRA87598.1; PID:gl1334478
A:Experimental source: female gametophytes
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC3
C:Keywords: mitochondrion

Query Match 88.9%; Score 32; DB 2; Length 73;
Best Local Similarity 83.3%; Pred. No. 26;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IFFWIK 6
Db 38 IFFWFK 43

RESULT 4
T32207
hypothetical protein T03D3.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T32207
R:Murray, J.; Wohldmann, P.; Bauer, C.; Biewald, T.
submitted to the EMBL Data Library, September 1997
A:Description: The sequence of C. elegans cosmid T03D3.
A:Reference number: T21136
A:Accession: T32207
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-374 <MUR>
A:Cross-references: EMBL:AF022980; PIDN:AAB69912.1; GSPDB:GN00023; CESP:T03D3.5
A:Experimental source: strain Bristol N2; clone T03D3
C:Genetics:
A:Gene: CESP:T03D3.5
A:Map position: 5
A:Introns: 17/3; 28/3; 43/2; 93/3; 155/3; 209/2; 244/2

Query Match 88.9%; Score 32; DB 2; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FFWIK 6
Db 36 FFWIK 40

RESULT 5
E90569

conserved hypothetical protein MYPU_4610 [imported] - Mycoplasma pulmonis (strain UAB
C:Species: Mycoplasma pulmonis
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C:Accession: E90569
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma p
A:Reference number: A99512; MUID:21267165; PMID:11353084
A:Accession: E90569
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-423 <KUR>
A:Cross-references: GB:AL445566; PID:gl4089875; PIDN:CAC13634.1; GSPDB:GN00153
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYPU_4610
A:Genetic code: SGC3

Query Match 88.9%; Score 32; DB 2; Length 423;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IFFWIK 6
Db 390 LFFWLK 395

RESULT 6
T51162
aminoglycoside acetyltransferase regulator [imported] - Providencia stuartii
C:Species: Providencia stuartii
C:Date: 28-Jul-2000 #sequence_revision 28-Jul-2000 #text_change 02-Sep-2000
C:Accession: T51162
R:Macinga, D.R.; Cook, G.M.; Poole, R.K.; Rather, P.N.
J. Bacteriol. 180, 128-135, 1998
A:Title: Identification and characterization of aarF, a locus required for production
artii.
A:Reference number: Z25318; MUID:98083065
A:Accession: T51162
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-544 <MAC>
A:Cross-references: EMBL:AF002165; PIDN:AAB96577.1
A:Experimental source: strain PR50
C:Genetics:
A:Gene: aarF
C:Superfamily: Synecocystis ABC transporter sir1919

Query Match 88.9%; Score 32; DB 2; Length 544;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FFWIK 6
Db 41 FFWIK 45

RESULT 7
AE3226
conserved hypothetical protein Atu5541 [imported] - Agrobacterium tumefaciens (strain
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AE3226
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; MCCI
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193

A:Accession: AE3226
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-660 <KUR>
A:Cross-references: GB:AE008687; PIDN:AL46227.1; PID:g17744003; GSPDB:GN00188
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu5541
A:Genome: plasmid

Query Match 88.9%; Score 32; DB 2; Length 660;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWIK 6
:|||||
Db 358 IFFWIE 363

RESULT 8

H71338
conserved hypothetical protein TP0325 - syphilis spirochete

C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C>Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 21-Jan-2000
C:Accession: H71338
R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwinn, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDevitt, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250; MUID:98332770
A:Accession: H71338
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-988 <COL>
A:Cross-references: GB:AE001212; GB:AE000520; MID:g3322597; PIDN:AAC65312.1; PID:g332260
A:Experimental source: strain Nichols
C:Genetics:
A:Gene: TP0325
C:Superfamily: syphilis spirochete conserved hypothetical protein TP0325

Query Match 88.9%; Score 32; DB 2; Length 988;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FFWIK 6
:|||||
Db 747 FFWIK 751

RESULT 9

AD0850
doubtful CDS found within S. typhi pathogenicity island [imported] - Salmonella enterica
C:Species: Salmonella enterica subsp. enterica serovar typhi
A:Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C:Accession: AD0850
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhimurium DT104
A:Reference number: AB0502; PMID:11677608
A:Accession: AD0850
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-86 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD05987.1; PID:g16503958; GSPDB:GN00176
C:Genetics:
A:Gene: STY3003

Query Match 86.1%; Score 31; DB 2; Length 86;
Best Local Similarity 66.7%; Pred. No. 47;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWIK 6
:|||||
Db 53 LFYWK 58

RESULT 10

D69381
hypothetical protein AF1052 - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: D69381
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.; Dodson, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirn, J.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.; Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes, Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaean Pyrococcus furiosus
A:Reference number: A69250; MUID:98049343
A:Accession: D69381
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-162 <NLE>
A:Cross-references: GB:AE001030; GB:AE000782; MID:g2689353; PIDN:AAB90193.1; PID:g264

Query Match 86.1%; Score 31; DB 2; Length 162;
Best Local Similarity 50.0%; Pred. No. 85;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWIK 6
:|||||
Db 68 VFYWK 73

RESULT 11

T33469
hypothetical protein F43B10.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 04-Mar-2000
C:Accession: T33469
R:Fulton, R.; Hawkins, J.; Rohlfing, T.
submitted to the EMBL Data Library, October 1998
A:Description: The sequence of C. elegans cosmid F43B10.
A:Reference number: Z21351
A:Accession: T33469
A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA
A:Residues: 1-247 <FUL>

A:Cross-references: EMBL:AF098500; PIDN:AAC67400.1; GSPDB:GN000028; CESP:F43B10.1
A:Experimental source: strain Bristol N2; clone F43B10
C:Genetics:
A:Gene: CESP:F43B10.1
A:Map position: X
A:Introns: 88/2; 103/3; 128/3
C:Superfamily: Caenorhabditis elegans hypothetical protein F43B10.1

Query Match 86.1%; Score 31; DB 2; Length 247;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWI 5
:|||||
Db 85 IFFWI 89

RESULT 12

AG2542
 hypothetical protein all7626 [imported] - Anabaena sp. (strain PCC 7120) plasmid pCC7120
 C:Species: Anabaena sp.
 A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC.7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
 C:Accession: AG2542
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena sp. strain PCC 7120
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AG2542
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-258 <KUR>
 A:Cross-references: GB:AP003602; PIDN:BAB77269.1; PID:g17134711; GSPDB:GN00181
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: all7626
 A:Genome: plasmid

Query Match

Best Local Similarity 86.1%; Score 31; DB 2; Length 258;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IFFWI 5

Db 20 IFFWI 24

RESULT 13

S75891
 hypothetical protein slr1174 - Synechocystis sp. (strain PCC 6803)
 C:Species: Synechocystis sp.
 A:Variety: PCC 6803
 C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
 C:Accession: S75891
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, M.
 DNA Res. 3, 109-136, 1996
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp.
 A:Reference number: S74322; MUID:97061201
 A:Accession: S75891
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-262 <KAN>
 A:Cross-references: EMBL:D90913; GB:AB001339; NID:gl653348; PIDN:BAA18350.1; PID:d101908
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match

Best Local Similarity 86.1%; Score 31; DB 2; Length 262;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IFFWI 5

Db 26 IFFWI 30

RESULT 14

A72238
 hemolysin - Thermotoga maritima (strain MSB8)
 C:Species: Thermotoga maritima
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
 C:Accession: A72238
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing

A:Reference number: A72200; MUID:99287316

A:Accession: A72238

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-267 <ARN>

A:Cross-references: GB:AE001802; GB:AE000512; NID:g4982133; PIDN:AAD36643.1; PID:g498

A:Experimental source: strain MSB8

C:Genetics:

A:Gene: TMI576

C:Superfamily: hemolysin homolog yqxC

Query Match

Best Local Similarity 86.1%; Score 31; DB 2; Length 267;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FFWIK 6

Db 237 FFWIK 241

RESULT 15

T05338
 hypothetical protein FIC12.211 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
 C:Accession: T05338
 R:Bevan, M.; Terry, N.; Ardiles, W.; Buysschaert, C.; Dasseville, R.; De Clerck, R.; Ewens, H.W.; Mayer, K.F.X.; Schueller, C.
 submitted to the Protein Sequence Database, April 1998
 A:Reference number: T05338
 A:Accession: T05338
 A:Molecule type: DNA
 A:Residues: 1-287 <BEV>
 A:Cross-references: EMBL:AL022224
 A:Experimental source: cultivar Columbia; BAC clone FIC12

C:Genetics:

A:Map position: 4

A:Introns: 37/2; 116/2; 243/3; 269/3

A:Note: FIC12.211

Query Match

Best Local Similarity 86.1%; Score 31; DB 2; Length 287;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FFWIK 6

Db 53 FFWIK 57

Search completed: August 6, 2002, 10:37:30

Job time: 47 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 10:40:50 ; Search time 18.57 Seconds
(without alignments)
12.510 Million cell updates/sec

Title: US-09-543-188A-23

Perfect score: 36

Sequence: 1 IFFWIK 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	32	88.9	544	1	UBIB_PROST
2	31	86.1	162	1	YA52_ARCFU
3	31	86.1	367	1	NUIC_MESVI
4	31	86.1	473	1	NORB_PEST
5	31	86.1	495	1	NU4M_MARPO
6	31	86.1	497	1	NU4M_ACACA
7	31	86.1	507	1	CPT7_RAT
8	31	86.1	639	1	FUUL_YEAST
9	31	86.1	669	1	GALC_CANFA
10	31	86.1	669	1	GALC_HUMAN
11	31	86.1	669	1	GALC_MACMU
12	31	86.1	684	1	ATY3_HUMAN
13	31	86.1	717	1	TRD1_ECOLI
14	31	86.1	738	1	TRD2_ECOLI
15	31	86.1	826	1	VILI_CHICK
16	31	86.1	951	1	SVV_ECOLI
17	31	86.1	955	1	SVV_BUCAL
18	31	86.1	1113	1	MGA2_YEAST
19	31	86.1	1256	1	FLU1_DROME
20	30	83.3	110	1	RT10_PEA
21	30	83.3	142	1	RADC_COXBU
22	30	83.3	349	1	YB64_SYNY3
23	30	83.3	367	1	NUIC_TOBAC
24	30	83.3	368	1	NUIC_MARPO
25	30	83.3	381	1	CYB_CARCH
26	30	83.3	381	1	CYB_ISUPA
27	30	83.3	411	1	LAG1_YEAST
28	30	83.3	470	1	MM12_HUMAN
29	30	83.3	537	1	AREH_SCHPO
30	30	83.3	642	1	YA2A_SCHPO
31	30	83.3	668	1	GALC_MOUSE
32	30	83.3	826	1	VILI_HUMAN
33	30	83.3	826	1	VILI_MOUSE

34 Q9x2d7 thermotoga
35 O51680 borrelia bu
36 Q9k8q8 bacillus ha
37 P11931 bacillus st
38 Q05873 bacillus su
39 Q55222 synechocyst
40 Q9jx22 neisseria m
41 Q9klh7 neisseria m
42 Q9kp73 vibrio chol
43 P43834 haemophilus
44 O83998 treponema p
45 Q9su58 arabisopsis

34 30 83.3 865 1 SYV_THEME
35 30 83.3 875 1 SYV_BORBU
36 30 83.3 880 1 SYV_BACHD
37 30 83.3 880 1 SYV_BACST
38 30 83.3 880 1 SYV_BACSU
39 30 83.3 910 1 SYV_SYNY3
40 30 83.3 945 1 SYV_NEIMA
41 30 83.3 945 1 SYV_NEIMB
42 30 83.3 953 1 SYV_VIBCH
43 30 83.3 954 1 SYV_HAEIN
44 30 83.3 956 1 SYV_TREPA
45 30 83.3 960 1 PMA4_ARATH

ALIGNMENTS

RESULT 1
ID UBIB_PROST
AC O07443;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable ubiquinone biosynthesis protein ubiB (Aminoglycoside
DE acetyltransferase regulator).
GN UBIB OR AARF.
OS Providencia stuartii.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Providencia.
OX NCBI_TaxID=588;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PR50;
RX MEDLINE=98083065; PubMed=9422602;
RA Macinga D.R., Cook G.M., Poole R.K., Rather P.N.;
RT "Identification and characterization of aarf, a locus required for
RT production of ubiquinone in Providencia stuartii and Escherichia coli
RT and for expression of 2'-N-acetyltransferase in P. stuartii.";
RL J. Bacteriol. 180:128-135(1998).
RN [2]
RP CHARACTERIZATION.
RC STRAIN=PR50;
RX MEDLINE=20416225; PubMed=10960098;
RA Poon W.W., Davis D.E., Ha H.T., Jonassen T., Rather P.N., Clarke C.F.;
RT "Identification of Escherichia coli ubiB, a gene required for the
RT first monooxygenase step in ubiquinone biosynthesis.";
RL J. Bacteriol. 182:5139-5146(2000).
CC -1- FUNCTION: REQUIRED, PROBABLY INDIRECTLY, FOR THE HYDROXYLATION OF
CC 2-OCTAPRENYLPHENOL TO 2-OCTAPRENYL-6-HYDROXY-PHENOL, THE FOURTH
CC STEP IN UBIQUINONE BIOSYNTHESIS. REQUIRED FOR THE EXPRESSION OF
CC 2'-N-ACETYLTRANSFERASE.
CC -1- SIMILARITY: BELONGS TO THE ABC1 FAMILY. UBIB SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/annouce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF002165; AAB96577.1;
DR InterPro: IPR004147; ABC1.
DR Pfam: PF03109; ABC1; 1.
KW Ubiquinone biosynthesis.
SQ SEQUENCE 544 AA; 62598 MW; E789A4E8185B4E96 CRC64;

Query Match 88.9%; Score 32; DB 1; Length 544;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RC STRAIN-ATCC 14405 / ZoBell;
 RA MEDLINE=98191362; PubMed=9521721;
 RX Cheesman M.R., Zumft W.G., Thomson A.J.;
 RT "The MCD and EPR of the heme centers of nitric oxide reductase
 from *Pseudomonas stutzeri*: evidence that the enzyme is
 structurally related to the heme-copper oxidases.";
 RL Biochemistry 37:3994-4000(1998).
 CC -!- FUNCTION: COMPONENT OF THE ANAEROBIC RESPIRATORY CHAIN THAT
 TRANSFORMS NITRATE TO DINITROGEN (DENITRIFICATION). NORB IS THE
 CATALYTIC SUBUNIT OF THE ENZYME COMPLEX. SHOWS PROTON PUMP
 ACTIVITY ACROSS THE MEMBRANE IN DENITRIFYING BACTERIAL CELLS. THE
 MONONITROGEN REDUCTION IS PROBABLY COUPLED TO ELECTRON TRANSPORT
 PHOSPHORYLATION.
 CC -!- CATALYTIC ACTIVITY: Nitrous oxide + acceptor + H(2)O = 2 nitric
 oxide + reduced acceptor.
 CC -!- SUBUNIT: HETERODIMER OF CYTOCHROMES B (LARGE SUBUNIT) AND C (SMALL
 SUBUNIT).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- INDUCTION: BY NITRIC OXIDE (PROBABLE).
 CC -!- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z28384; CAA82229.1; -
 DR InterPro; IPR000883; COX1.
 DR PROSITE; PS00077; COX1; 1.
 KW Oxidoreductase; Heme; Iron; Transmembrane; Respiratory chain.
 FT INIT_MET 0 0
 FT METAL 59 59 IRON (LOW-SPIN HEME) (PROBABLE).
 FT METAL 206 206 IRON B (PROBABLE).
 FT METAL 257 257 IRON B (PROBABLE).
 FT METAL 258 258 IRON B (PROBABLE).
 FT METAL 346 346 IRON (HIGH-SPIN HEME) (PROBABLE).
 FT METAL 348 348 IRON (LOW-SPIN HEME) (PROBABLE).
 FT TRANSMEM 18 38 POTENTIAL.
 FT TRANSMEM 60 80 POTENTIAL.
 FT TRANSMEM 94 114 POTENTIAL.
 FT TRANSMEM 144 164 POTENTIAL.
 FT TRANSMEM 168 188 POTENTIAL.
 FT TRANSMEM 206 226 POTENTIAL.
 FT TRANSMEM 242 262 POTENTIAL.
 FT TRANSMEM 269 289 POTENTIAL.
 FT TRANSMEM 307 327 POTENTIAL.
 FT TRANSMEM 347 367 POTENTIAL.
 FT TRANSMEM 389 409 POTENTIAL.
 FT TRANSMEM 432 452 POTENTIAL.
 SQ SEQUENCE 473 AA; 53006 MW; 39E999DC1DD7F757; CRC64;
 Query Match 86.1%; Score 31; DB 1; Length 473;
 Best Local Similarity 66.7%; Pred. No. 98;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IFFWIK 6
 Db 434 IFFWLIR 439
 RESULT 5
 NU4M_MARPO STANDARD; PRT; 495 AA;
 ID NU4M_MARPO
 AC P26848;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3).
 GN ND4 OR NAD4.

OS Marchantia polymorpha (Liverwort).
 OG Mitochondrion.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Marchantiophyta;
 OC Marchantiales; Marchantiaceae; Marchantia.
 OX NCBI_TaxID=3197;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92114051; PubMed=1731062;
 RA Oda K., Yamato K., Ohta E., Nakamura Y., Takemura M., Nozato N.,
 RA Akashi K., Kanegae T., Ogura Y., Kohchi T., Ohshima K.;
 RT "Gene organization deduced from the complete sequence of liverwort
 Marchantia polymorpha mitochondrial DNA. A primitive form of plant
 mitochondrial genome.";
 RL J. Mol. Biol. 223:1-7(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93247547; PubMed=8483448;
 RA Nozato N., Oda K., Yamato K., Ohta E., Takemura M., Akashi K.,
 RA Fukuzawa H., Ohshima K.;
 RT "Cloning and characterization of the complete sequence of the
 NADH dehydrogenase, nad5, nad4, nad2, in Marchantia polymorpha.";
 RL Mol. Gen. Genet. 237:343-350(1993).
 CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M68929; AAC09398.1; -
 DR PIR; S25942; S25942.
 DR Mendel; 2061; MARPO:nad4;1.
 DR InterPro; IPR003918; NADHub_oxdrdctse4.
 DR InterPro; IPR001750; Oxidored_g1.
 DR Pfam; PF00361; oxidored_g1; 1.
 DR PRINTS; PR01437; NUOXDRDTASE4.
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
 SQ SEQUENCE 495 AA; 56311 MW; 0F75894D6CAAAED4 CRC64;
 Query Match 86.1%; Score 31; DB 1; Length 495;
 Best Local Similarity 66.7%; Pred. No. 1e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IFFWIK 6
 Db 53 LFFWIR 58
 RESULT 6
 NU4M_ACACA STANDARD; PRT; 497 AA.
 ID NU4M_ACACA
 AC Q37375;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3).
 GN ND4 OR NAD4.
 OS Acanthamoeba castellanii (Amoeba).
 OG Mitochondrion.
 OC Eukaryota; Acanthamoebidae; Acanthamoeba.
 OX NCBI_TaxID=5755;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN-ATCC 30010 / NEFF;
 RA MEDLINE=95147275; PubMed=7844823;
 RA Burger G., Plante I., Loneragan K.M., Gray M.W.;
 RT "The mitochondrial DNA of the amoeboid protozoan, Acanthamoeba
 castellanii: complete sequence, gene content and genome
 organization.";
 RT organization.";

RL J. Mol. Biol. 245:522-537(1995).
CC -|- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U12386; AAD11826.1; -
DR InterPro: IPR003918; NADH_oxdrdctse4.
DR InterPro: IPR001750; Oxidored_g1.
DR Pfam: PF00361; oxidored_g1; 1.
DR PRINTS: PR01437; NUOXDRDTASE4.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 497 AA; 57768 MW; 9C0B637E65C0F3B5 CRC64;

Query Match 86.1%; Score 31; DB 1; Length 497;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWI 5
DB 53 IFFWI 57

RESULT 7
CPT7_RAT
ID CPT7_RAT STANDARD; PRT; 507 AA.
AC P11715;
DT 01-OCT-1989 (Rel. 12; Created)
DT 01-AUG-1990 (Rel. 15; Last sequence update)
DT 15-JUL-1999 (Rel. 38; Last annotation update)
DE Cytochrome P450 17 (EC 1.14.99.9) (CYPXVII) (P450-C17) (Steroid
DE 17-alpha-Hydroxylase/17,20 lyase).
GN CYP17.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89295447; PubMed=2786990;
RA Fevold H.R., Lorence M.C., McCarthy J.L., Trant J.M., Kagimoto M.,
RA Waterman M.R., Mason J.I.;
RT "Rat P450(17 alpha) from testis: characterization of a full-length
RT cDNA encoding a unique steroid hydroxylase capable of catalyzing both
RT delta 4- and delta 5-steroid-17,20-lyase reactions.";
RL Mol. Endocrinol. 3:968-975(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=89076306; PubMed=3264499;
RA Namiki M., Kitamura M., Buczko E., Dufau M.L.;
RT "Rat testis P-450(17)alpha cDNA: the deduced amino acid sequence,
RT expression and secondary structural configuration.";
RL Biochem. Biophys. Res. Commun. 157:705-712(1988).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RX MEDLINE=95217329; PubMed=7702752;
RA Givens C.R., Zhang P., Bair S.R., Mellon S.H.;
RT "Transcriptional regulation of rat cytochrome P450c17 expression in
RT mouse Leydig MA-10 and adrenal Y-1 cells: identification of a single
RT protein that mediates both basal and CAMP-induced activities.";
RL DNA Cell Biol. 13:1087-1098(1994).
RN [4]
RP SEQUENCE OF 271-507 FROM N.A.
RX MEDLINE=88280759; PubMed=3260774;
RA Nishihara M., Winters C.A., Buzko E., Waterman M.R., Dufau M.L.;

RT "Hormonal regulation of rat Leydig cell cytochrome P-45017 alpha mRNA
RT levels and characterization of a partial length rat P-45017 alpha
RT cDNA.";
RL Biochem. Biophys. Res. Commun. 154:151-158(1988).
RN [5]
RP SEQUENCE OF 273-507 FROM N.A.
RX MEDLINE=90046678; PubMed=2554289;
RA Mellon S.H., Vaisse C.;
RT "CAMP regulates P450sc gene expression by a
RT cycloheximide-insensitive mechanism in cultured mouse Leydig MA-10
RT cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:7775-7779(1989).
CC -|- FUNCTION: CONVERSION OF PREGNENOLONE AND PROGESTERONE TO THEIR
CC 17-ALPHA-HYDROXYLATED PRODUCTS AND SUBSEQUENTLY TO
CC DEHYDROEPIANDROSTERONE (DHEA) AND ANDROSTENEDIONE. CATALYZES
CC BOTH THE 17-ALPHA-HYDROXYLATION AND THE 17,20-LYASE REACTION.
CC INVOLVED IN SEXUAL DEVELOPMENT DURING FETAL LIFE AND AT PUBERTY.
CC -|- CATALYTIC ACTIVITY: A steroid + AH(2) + O(2) -> a 17-alpha-
CC hydroxysteroid + A + H(2)O.
CC -|- ENZYME REGULATION: REGULATED PREDOMINANTLY BY INTRACELLULAR CAMP
CC LEVELS.
CC -|- PATHWAY: KEY ENZYME IN STEROIDOGENIC PATHWAY.
CC -|- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14086; CAA32248.1; -
DR EMBL; X69816; CAA49470.1; -
DR EMBL; M31681; AAA41777.1; -
DR EMBL; M22204; AAA41783.1; -
DR EMBL; M21208; AAA41050.1; -
DR EMBL; M27282; AAA41779.1; -
DR PIR; A27659; A27659.
DR PIR; A30828; A30828.
DR PIR; S16719; S16719.
DR InterPro: IPR001128; Cyt_P450.
DR Pfam: PF00067; P450; 1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; 1.
KW Electron transport; Oxidoreductase; Monooxygenase; Membrane;
KW Heme; Steroidogenesis.
FT BINDING 441 441 HEME (BY SIMILARITY).
FT CONFLICT 505 506 VS -> LT (IN REF. 4).
SQ SEQUENCE 507 AA; 57250 MW; A535600F7E6A399B CRC64;

Query Match 86.1%; Score 31; DB 1; Length 507;
Best Local Similarity 80.0%; Pred. No. 1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 FFWIK 6
DB 15 FFWIK 19

RESULT 8
FUI1_YEAST
ID FUI1_YEAST STANDARD; PRT; 639 AA.
AC P38196;
DT 01-OCT-1994 (Rel. 30; Created)
DT 01-OCT-1994 (Rel. 30; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Uridine permease.
RN FUI1 OR YBL042C OR YBL0406.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=95176707; PubMed=7871888;
RA de Wergifosse P., Jacques B., Joniaux J.-L., Purnelle B., Skala J.,
RA Goffeau A.;
RT "The sequence of a 22.4 kb DNA fragment from the left arm of yeast
RT chromosome II reveals homologues to bacterial proline synthetase and
RT murine alpha-adaptin, as well as a new permease and a DNA-binding
RT protein.";
RL Yeast 10:1489-1496(1994).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=98146544; PubMed=9485596;
RA Wagner R., de Montigny J., de Wergifosse P., Souciet J.-L., Potier S.;
RT "The ORF YBL042 of *Saccharomyces cerevisiae* encodes a uridine
RT permease.";
RL FEMS Microbiol. Lett. 159:69-75(1998).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=20408929; PubMed=10827169;
RA Vickers M.F., Yeo S.Y., Baldwin S.A., Young J.D., Cass C.E.;
RT "Nucleoside transporter proteins of *Saccharomyces cerevisiae*.
RT Demonstration of a transporter (FUI1) with high uridine selectivity
RT in plasma membranes and a transporter (FUN26) with broad nucleoside
RT selectivity in intracellular membranes.";
RL J. Biol. Chem. 275:25931-25938(2000).
CC -!- FUNCTION: HIGH-AFFINITY TRANSPORT OF URIDINE.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE ALLANTOIN PERMEASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X78214; CA55059.1; -
DR EMBL; Z35803; CA84862.1; -
DR PIR; S45776; S45776.
DR SGD; S0000138; FUI1.
DR InterPro; IPR001248; Transp_cyt_pur.
DR Pfam; PF02133; Transp_cyt_pur; 1.
KW Transmembrane; Transport.
FT TRANSMEM 163 180 POTENTIAL.
FT TRANSMEM 201 225 POTENTIAL.
FT TRANSMEM 260 276 POTENTIAL.
FT TRANSMEM 284 305 POTENTIAL.
FT TRANSMEM 368 392 POTENTIAL.
FT TRANSMEM 417 435 POTENTIAL.
FT TRANSMEM 461 477 POTENTIAL.
FT TRANSMEM 484 507 POTENTIAL.
FT TRANSMEM 538 562 POTENTIAL.
FT TRANSMEM 573 590 POTENTIAL.
SQ SEQUENCE 639 AA; 72165 MW; 757325DEF4100399 CRC64;

Query Match 86.1%; Score 31; DB 1; Length 639;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IFFWI 5
Db 133 IFFWI 137

RESULT 9
GALC_CANFA
ID GALC_CANFA STANDARD; PRT; 669 AA.
AC P54804;

DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Galactocerebrosidase precursor (EC 3.2.1.46) (GALCERASE)
DE (Galactosylceramidase) (Galactosylceramide beta-galactosidase)
DE (Galactocerebrosidase) (Galactosylceramide beta-galactosidase)
GN GALC.
OS *Canis familiaris* (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANTS GLD SER-158 AND SER-639.
RX MEDLINE=96299640; PubMed=8661004;
RA Victoria T., Rafi M.A., Wenger D.A.;
RT "Cloning of the canine GALC cDNA and identification of the mutation
RT causing globoid cell leukodystrophy in West Highland White and Cairn
RT terriers.";
RL Genomics 33:457-462(1996).
CC -!- FUNCTION: HYDROLYSES THE GALACTOSE ESTER BONDS OF
CC GALACTOSYLKERAMIDE, GALACTOSYLSPHINGOSINE, LACTOSYLKERAMIDE, AND
CC MONOGALACTOSYLDIGLYCERIDE. ENZYME WITH VERY LOW ACTIVITY
CC RESPONSIBLE FOR THE LYSOSOMAL CATABOLISM OF GALACTOSYLKERAMIDE, A
CC MAJOR LIPID IN MYELIN, KIDNEY AND EPITHELIAL CELLS OF SMALL
CC INTESTINE AND COLON (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: D-galactosyl-N-acylsphingosine + H(2)O = D-
CC galactose + N-acylsphingosine.
CC -!- SUBCELLULAR LOCATION: Lysosomal (By similarity).
CC -!- DISEASE: DEFECTS IN GALC ARE THE CAUSE OF GLOBOID CELL
CC LEUKODYSTROPHY (GLD). THIS DEFICIENCY RESULTS IN THE INSUFFICIENT
CC CATABOLISM OF SEVERAL GALACTOLIPIDS THAT ARE IMPORTANT IN THE
CC PRODUCTION OF NORMAL MYELIN.
CC -!- SIMILARITY: BELONGS TO FAMILY 59 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; L76184; AAB37752.1; -
DR InterPro; IPR001286; Glyco_hydro_59.
DR Pfam; PF02057; Glyco_hydro_59; 1.
DR PRINTS; PR00850; GLHYDRLASE59.
KW Hydrolase; Glycosidase; Glycoprotein; Signal; Disease mutation.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 669 GALACTOCEREBROSIDASE.
FT CARBOHYD 363 363 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 387 387 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 543 543 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 586 586 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 158 158 Y -> S (IN GLD).
FT VARIANT 639 639 P -> S (IN GLD).
SQ SEQUENCE 669 AA; 75316 MW; 60E298B024EE154C CRC64;

Query Match 86.1%; Score 31; DB 1; Length 669;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IFFWI 5
Db 579 IFFWI 583

RESULT 10
GALC_HUMAN
ID GALC_HUMAN STANDARD; PRT; 669 AA.
AC P54803;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)

16-OCT-2001 (Rel. 40, Last annotation update)
DE Galactocerebroside precursor (EC 3.2.1.46) (GALCERASE)
DE (Galactosylceramide) (Galactosylceramide beta-galactosidase)
DE (Galactocerebroside beta-galactosidase).
GN GALC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE-Placenta, and Skin fibroblast;
RX MEDLINE=94128088; PubMed=8297359;
RA Sakai N., Inui K., Fujii N., Fukushima H., Nishimoto J.,
RA Yanagihara I., Isegawa Y., Iwamatsu A., Okada S.;
RT "Krabbe disease: isolation and characterization of a full-length cDNA
RT for human galactocerebroside.";
RL Biochem. Biophys. Res. Commun. 198:485-491(1994).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-59 AND 436-454.
RC TISSUE-Brain, and Testis;
RX MEDLINE=94108435; PubMed=8281145;
RA Chen Y.Q., Rafi M.A., de Gala G., Wenger D.A.;
RT "Cloning and expression of cDNA encoding human galactocerebroside,
RT the enzyme deficient in globoid cell leukodystrophy.";
RL Hum. Mol. Genet. 2:1841-1845(1993).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=95324938; PubMed=7601472;
RA Luzzi P., Rafi M.A., Wenger D.A.;
RT "Structure and organization of the human galactocerebroside (GALC)
RT gene.";
RL Genomics 26:407-409(1995).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=98094242; PubMed=9434153;
RA Sakai N., Fukushima H., Inui K., Fu L., Nishigaki T., Yanagihara I.,
RA Tatsumi N., Ozono K., Okada S.;
RT "Human galactocerebroside gene: promoter analysis of the 5'-flanking
RT region and structural organization.";
RL Biochim. Biophys. Acta 1395:62-67(1998).
RN [5]
RP SEQUENCE OF 27-45 AND 436-454, AND CHARACTERIZATION.
RC TISSUE-Urine;
RX MEDLINE=94002192; PubMed=8399327;
RA Chen Y.Q., Wenger D.A.;
RT "Galactocerebroside from human urine: purification and partial
RT characterization.";
RL Biochim. Biophys. Acta 1170:53-61(1993).
RN [6]
RP REVIEW ON GLD MUTATIONS.
RX MEDLINE=97478285; PubMed=9338580;
RA Wenger D.A., Rafi M.A., Luzzi P.;
RT "Molecular genetics of Krabbe disease (globoid cell leukodystrophy):
RT diagnostic and clinical implications.";
RL Hum. Mutat. 10:268-279(1997).
RN [7]
RP VARIANTS GLD ALA-302 AND GLY-550.
RX MEDLINE=96121583; PubMed=8595408;
RA Tatsumi N., Inui K., Sakai N., Fukushima H., Nishimoto J.,
RA Yanagihara I., Nishigaki T., Tsukamoto H., Fu L., Tanike M.,
RA Okada S.;
RT "Molecular defects in Krabbe disease.";
RL Hum. Mol. Genet. 4:1865-1868(1995).
RN [8]
RP VARIANTS GLD ASN-528 AND SER-583.
RX MEDLINE=96198195; PubMed=8786069;
RA Rafi M.A., Luzzi P., Zlotogora J., Wenger D.A.;
RT "Two different mutations are responsible for Krabbe disease in the
RT Arab and Moslem Arab populations in Israel.";
RL Hum. Genet. 97:304-308(1996).
CC -1- FUNCTION: HYDROLYSES THE GALACTOSE ESTER BONDS OF
CC GALACTOSYLCEAMIDE, GALACTOSYLSPHINGOSINE, LACTOSYLCEAMIDE, AND

MONOGALACTOSYLDIGLYCERIDE. ENZYME WITH VERY LOW ACTIVITY
RESPONSIBLE FOR THE LYSOSOMAL CATABOLISM OF GALACTOSYLCEAMIDE, A
MAJOR LIPID IN MYELIN, KIDNEY AND EPITHELIAL CELLS OF SMALL
INTESTINE AND COLON. HAS AN OPTIMAL PH BETWEEN 4.0 AND 4.4.
ACTIVITY IS LOST WHEN HEATED AT 52 DEGREES CELSIUS FOR FIVE
MINUTES.
-1- CATALYTIC ACTIVITY: D-galactosyl-N-acylsphingosine + H(2)O - D-
galactose + N-acylsphingosine.
-1- SUBCELLULAR LOCATION: Lysosomal.
-1- ALTERNATIVE PRODUCTS: TWO FORMS ARE PRODUCED BY ALTERNATIVE
SPLICING. THE SEQUENCE SHOWN HERE IS THAT OF THE LONG FORM.
-1- TISSUE SPECIFICITY: HIGHEST LEVEL OF ACTIVITY IN TESTES COMPARED
TO BRAIN, KIDNEY, PLACENTA AND LIVER. CAN ALSO BE FOUND IN URINE.
-1- DISEASE: DEFECTS IN GALC ARE THE CAUSE OF GLOBOID CELL
LEUKODYSTROPHY (GLD) (OR KRABBE DISEASE). THIS AUTOSOMAL RECESSIVE
DISORDER DEFICIENCY RESULTS IN THE INSUFFICIENT CATABOLISM OF
SEVERAL GALACTOLIPIDS THAT ARE IMPORTANT IN THE PRODUCTION OF
NORMAL MYELIN. CLINICALLY THE MOST FREQUENT FORM IS THE INFANTILE
FORM. MOST PATIENTS (90%) PRESENT BEFORE SIX MONTHS OF AGE WITH
IRRITABILITY, SPASTICITY, ARREST OF MOTOR AND MENTAL DEVELOPMENT,
AND BOUTS OF TEMPERATURE ELEVATION WITHOUT INFECTION. THIS IS
FOLLOWED BY MYOCLONIC JERKS OF ARMS AND LEGS, OPOSTHOTONUS,
HYPERTONIC FITS, AND MENTAL REGRESSION, WHICH PROGRESSES TO A
SEVERE DECEREBRATE CONDITION WITH NO VOLUNTARY MOVEMENTS AND DEATH
FROM RESPIRATORY INFECTIONS OR CEREBRAL HYPERTENSIA BEFORE 2
YEARS OF AGE. HOWEVER, A SIGNIFICANT NUMBER OF CASES WITH LATER
ONSET, PRESENTING WITH UNEXPLAINED BLINDNESS, WEAKNESS AND/OR
PROGRESSIVE MOTOR, AND SENSORY NEUROPATHY THAT CAN PROGRESS TO
SEVERE MENTAL INCAPACITY AND DEATH, HAVE BEEN IDENTIFIED.
-1- SIMILARITY: BELONGS TO FAMILY 59 OF GLYCOSYL HYDROLASES.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).

EMBL; D25283; BAA04971.1; -
EMBL; D25284; BAA04972.1; -
EMBL; L23116; AAA16645.1; -
EMBL; L38559; AAA80975.1; JOINED.
EMBL; L38545; AAA80975.1; JOINED.
EMBL; L38546; AAA80975.1; JOINED.
EMBL; L38547; AAA80975.1; JOINED.
EMBL; L38548; AAA80975.1; JOINED.
EMBL; L38549; AAA80975.1; JOINED.
EMBL; L38550; AAA80975.1; JOINED.
EMBL; L38551; AAA80975.1; JOINED.
EMBL; L38552; AAA80975.1; JOINED.
EMBL; L38553; AAA80975.1; JOINED.
EMBL; L38555; AAA80975.1; JOINED.
EMBL; L38556; AAA80975.1; JOINED.
EMBL; L38557; AAA80975.1; JOINED.
EMBL; L38558; AAA80975.1; JOINED.
EMBL; D81181; BAA24902.1; -
EMBL; D84232; BAA24902.1; JOINED.
EMBL; D84233; BAA24902.1; JOINED.
EMBL; D84234; BAA24902.1; JOINED.
EMBL; D84266; BAA24902.1; JOINED.
EMBL; D84267; BAA24902.1; JOINED.
EMBL; D84268; BAA24902.1; JOINED.
EMBL; D84269; BAA24902.1; JOINED.
EMBL; D84363; BAA24902.1; JOINED.
EMBL; D84364; BAA24902.1; JOINED.
EMBL; D84365; BAA24902.1; JOINED.
EMBL; D84366; BAA24902.1; JOINED.
EMBL; D84367; BAA24902.1; JOINED.
EMBL; D84368; BAA24902.1; JOINED.
EMBL; D84369; BAA24902.1; JOINED.
EMBL; D84393; BAA24902.1; JOINED.


```

DR DR 245200;
DR InterPro: IPR001286; Glyco_hydro_59.
DR Pfam: PF02057; Glyco_hydro_59; 1.
DR PRINTS: PR00850; GLHYDLASE59.
KW Hydrolase; Glycosidase; Glycoprotein; Signal; Alternative splicing;
KW Disase mutation; Polymorphism.
FT SIGNAL 1 26
FT CHAIN 27 669
FT CARBOHYD 127 127
FT CARBOHYD 363 363
FT CARBOHYD 387 387
FT CARBOHYD 540 540
FT CARBOHYD 543 543
FT CARBOHYD 586 586
FT VARSPLIC 372 391
FT VARSPLIC 392 669
FT VARIANT 95 95
FT VARIANT 95 95
FT VARIANT 96 96
FT VARIANT 101 101
FT VARIANT 171 171
FT VARIANT 178 178
FT VARIANT 232 232
FT VARIANT 234 234
FT VARIANT 247 247
FT VARIANT 268 268
FT VARIANT 270 270
FT VARIANT 279 279
FT VARIANT 287 287
FT VARIANT 302 302
FT VARIANT 380 380
FT VARIANT 384 384

Query Match 86.1%; Score 31; DB 1; Length 669;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IFFWI 5
Db 579 IFFWI 583

RESULT 11
ID GALC_MACMU STANDARD; PRT; 669 AA.
AC 002791;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE Galactocerebrosidase precursor (EC 3.2.1.46) (GALCERASE)
DE (Galactosylceramide) (Galactosylceramide beta-galactosidase)
DE (Galactocerebroside beta-galactosidase).
GN GALC.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97336058; PubMed=9192853;
RT Luzi P., Rafi M.A., Victoria T., Baskin G.B., Wenger D.A.;
RA "Characterization of the rhesus monkey galactocerebrosidase (GALC)
RT cDNA and gene identification of the mutation causing globoid cell
RT leukodystrophy (Krabbe disease) in this primate.";
RL Genomics 42:319-324(1997).
CC -1- FUNCTION: HYDROLYSES THE GALACTOSE ESTER BONDS OF
CC GALACTOSYLCEAMIDE, GALACTOSYLSPHINGOSINE, LACTOSYLCEAMIDE, AND
CC MONOACETOSYLDIGLYCERIDE. ENZYME WITH VERY LOW ACTIVITY
CC RESPONSIBLE FOR THE LYSOSOMAL CATABOLISM OF GALACTOSYLCEAMIDE, A
CC MAJOR LIPID IN MYELIN. KIDNEY AND EPITHELIAL CELLS OF SMALL
CC INTESTINE AND COLON (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: D-galactosyl-N-acylsphingosine + H(2)O -> D-
CC galactose + N-acylsphingosine.
CC -1- SUBCELLULAR LOCATION: Lysosomal.
CC -1- DISEASE: DEFECTS IN GALC ARE THE CAUSE OF GLOBOID CELL
CC LEUKODYSTROPHY (GLD) (OR KRABBE DISEASE). THIS DEFICIENCY RESULTS
CC IN THE INSUFFICIENT CATABOLISM OF SEVERAL GALACTOLIPIDS THAT ARE
CC IMPORTANT IN THE PRODUCTION OF NORMAL MYELIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 59 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U87477; AAB58575.1;
DR EMBL; U87462; AAB58575.1; JOINED.
DR EMBL; U87463; AAB58575.1; JOINED.
DR EMBL; U87464; AAB58575.1; JOINED.
DR EMBL; U87465; AAB58575.1; JOINED.
DR EMBL; U87466; AAB58575.1; JOINED.
DR EMBL; U87467; AAB58575.1; JOINED.
DR EMBL; U87468; AAB58575.1; JOINED.
DR EMBL; U87469; AAB58575.1; JOINED.
DR EMBL; U87470; AAB58575.1; JOINED.
DR EMBL; U87471; AAB58575.1; JOINED.
DR EMBL; U87472; AAB58575.1; JOINED.
DR EMBL; U87473; AAB58575.1; JOINED.
DR EMBL; U87474; AAB58575.1; JOINED.
DR EMBL; U87475; AAB58575.1; JOINED.
DR EMBL; U87476; AAB58575.1; JOINED.
DR EMBL; U87628; AAB58576.1;
DR InterPro: IPR001286; Glyco_hydro_59.
DR Pfam: PF02057; Glyco_hydro_59; 1.
DR PRINTS: PR00850; GLHYDLASE59.
KW Hydrolase; Glycosidase; Glycoprotein; Signal.
FT SIGNAL 1 26
FT CHAIN 27 668
FT CARBOHYD 127 127
FT CARBOHYD 363 363
FT CARBOHYD 387 387
FT CARBOHYD 435 435
FT CARBOHYD 540 540
FT CARBOHYD 543 543
FT CARBOHYD 586 586
FT SEQUENCE 669 AA; 75206 MW; F09169EDC2B66C07 CRC64;

Query Match 86.1%; Score 31; DB 1; Length 669;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IFFWI 5

```


This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; M29254; AAC83928.1; -;
EMBL; U01159; AAC44181.1; -;
EMBL; AP001918; BAA97972.1; -;
EMBL; X06915; CAA30013.1; ALT_SEQ.
EMBL; M54796; AAA98083.1; -;
EMBL; X57431; CAA40678.1; -;
EMBL; X57428; CAA40674.1; -;
PIR; JS0293; BVFCAD.
PIR; S01758; S01758.
EcGene; EG40117; trad.
InterPro; IPR003688; TRAG.
PIR; PF02534; TRAG; 1.
Plasmid; Inner membrane; Conjugation; Transmembrane; ATP-binding;
Complete proteome.
DOMAIN 1 27 CYTOPLASMIC (PROBABLE).
TRANSMEM 28 47 PROBABLE.
DOMAIN 48 104 PERIPLASMIC (PROBABLE).
TRANSMEM 105 130 PROBABLE.
DOMAIN 131 717 CYTOPLASMIC (PROBABLE).
NP_BIND 192 199 ATP (POTENTIAL).
CONFLICT 19 47 IRMSQIANIMLYCLFIFFWILYGLVLMV -> YPHVQPR
CONFLICT 19 34 QYHALPVEIFLDRWGFMD (IN REF. 1 AND 6).
CONFLICT 19 34 IRMSQIANIMLYCLF -> YPHVQPRQYHALLPV
(IN REF. 4).
CONFLICT 52 54 QTF -> TDV (IN REF. 1 AND 6).
CONFLICT 58 62 CIYWM -> LFTV (IN REF. 1 AND 6).
CONFLICT 190 190 R -> L (IN REF. 3).
CONFLICT 269 270 RD -> PM (IN REF. 3).
CONFLICT 321 325 VIHRQ -> RNSPA (IN REF. 3).
SEQUENCE 717 AA; 81683 MW; F4B564EDD90EB914 CRC64;

Query Match 86.1%; Score 31; DB 1; Length 717;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWI 5
Db 35 IFFWI 39

RESULT 14
TRD2_ECOLI
ID TRD2_ECOLI STANDARD; PRT; 738 AA.
AC P22708; 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Trad protein.
GN Trad.
OS Escherichia coli.
OG Plasmid IncFII R100.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90317835; PubMed=2164585;
RA Yoshioaka Y., Fujita Y., Ohtsubo E.;
RT "Nucleotide sequence of the promoter-distal region of the tra operon
of plasmid R100, including traI (DNA helicase I) and traD genes.";
J. Mol. Biol. 214:39-53(1990).
CC -1- FUNCTION: TRAD IS A CELL ENVELOPE PROTEIN. THE FUNCTION OF TRAD
IS UNKNOWN BUT IT IS ESSENTIAL FOR DNA TRANSFER. IT MAY FORM OR
MAY BE PART OF A NON-SPECIFIC PORE FOR DNA EXPORT, OR IT MAY

DIRECTLY ENERGIZES DNA TRANSPORT.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -1- SIMILARITY: TO THE TRAD OF PLASMID F.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

EMBL; X55815; CAA39336.1; -;
PIR; S10659; S10659.
InterPro; IPR003688; TRAG.
PIR; PF02534; TRAG; 1.
Plasmid; Inner membrane; Conjugation; Transmembrane; ATP-binding;
Repeat.
KW Repeat.
FT TRANSMEM 28 44 POTENTIAL.
FT TRANSMEM 113 129 POTENTIAL.
FT TRANSMEM 396 413 POTENTIAL.
FT NP_BIND 192 199 ATP (POTENTIAL).
FT DOMAIN 617 646 10 X 3 AA TANDEM REPEATS OF Q-Q-P.
SQ SEQUENCE 738 AA; 83900 MW; 84CB1F48245E766F CRC64;

Query Match 86.1%; Score 31; DB 1; Length 738;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWI 5
Db 35 IFFWI 39

RESULT 15
VILI_CHICK
ID VILI_CHICK STANDARD; PRT; 826 AA.
AC P02640;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Villin.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88276884; PubMed=2839826;
RA Bazari W.L., Matsudaira P., Wallik M., Smeal T., Jakes R., Ahmed Y.;
RT "Villin sequence and peptide map identify six homologous domains.";
Proc. Natl. Acad. Sci. U.S.A. 85:4986-4990(1988).
RN [2]
RP SEQUENCE OF 751-326.
RX MEDLINE=81264203; PubMed=6790532;
RA Glenney J.R. Jr., Geisler N., Kaulfus P., Weber K.;
RT "Demonstration of at least two different actin-binding sites in
villin, a calcium-regulated modulator of F-actin organization.";
J. Biol. Chem. 256:8156-8161(1981).
RN [3]
RP CALCIUM-BINDING SITES.
RX MEDLINE=83082892; PubMed=6848508;
RA Hesterberg L.K., Weber K.;
RT "Demonstration of three distinct calcium-binding sites in villin, a
modulator of actin assembly.";
J. Biol. Chem. 258:365-369(1983).
RN [4]
RP STRUCTURE BY NMR OF 1-127.
RX MEDLINE=94191534; PubMed=8142900;
RA Markus M.A., Nakayama T., Matsudaira P., Wagner G.;
RT "Solution structure of villin 147, a domain conserved among actin-

severing proteins.";
RL Protein Sci. 3:70-81(1994).
RN [5]
RP STRUCTURE BY NMR OF 1-127.
RX MEDLINE-97337440; PubMed-9194180;
RA Markus M.A., Matsudaira P., Wagner G.;
RT "Refined structure of villin 147 and a detailed comparison with other
actin-severing domains".
RN Protein Sci. 6:1197-1209(1997).
RN [6]
RP STRUCTURE BY NMR OF 792-826.
RX MEDLINE-97307248; PubMed-9164455;
RA McKnight C.J., Matsudaira P.T., Kim P.S.;
RT "NMR structure of the 35-residue villin headpiece subdomain".
RN Nat. Struct. Biol. 4:180-184(1997).
CC -!- FUNCTION: CA(2+)-REGULATED ACTIN-BINDING PROTEIN. ITS CAPPING
FUNCTION IS LOCALIZED TO DOMAIN I; ITS SEVERING ACTION IS A
COOPERATION BETWEEN AN F-ACTIN BINDING SITE OF DOMAIN II AND THE
CAPPING ACTIVITY OF DOMAIN I.
CC -!- SUBUNIT: MONOMER.
CC -!- TISSUE SPECIFICITY: COMPONENT OF BRUSH BORDER-MICROVILLI.
CC -!- SIMILARITY: BELONGS TO THE VILLIN/GELSOLIN FAMILY.
CC -!- SIMILARITY: CONTAINS 6 GELSOLIN-LIKE REPEATS.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; J03781; AAA49133.1; -.
DR PIR; A03082; A03082.
DR PIR; A31822; A31822.
DR PDB; 2VIK; 01-APR-97.
DR PDB; 2VIL; 01-APR-97.
DR PDB; 1VII; 12-AUG-97.
DR InterPro; IPR001974; Gelsolin.
DR InterPro; IPR003128; VHP.
DR Pfam; PF00626; Gelsolin; 6.
DR Pfam; PF02209; VHP; 1.
DR PRINTS; PR00597; GELSOLIN.
DR SMART; SM00262; GEL; 6.
DR SMART; SM00153; VHP; 1.
KW Cytoskeleton; Calcium; Actin-binding; Capping protein; Repeat;
KW 3D-structure.
FT DOMAIN 1 734
FT DOMAIN 735 826
FT REPEAT 27 76
FT REPEAT 148 188
FT REPEAT 265 309
FT REPEAT 408 457
FT REPEAT 528 568
FT REPEAT 631 672
FT SITE 820 823
FT SITE 129 137
FT SITE 112 119
FT SITE 138 146
FT STRAND 21 23
FT STRAND 30 32
FT TURN 34 36
FT STRAND 39 41
FT STRAND 44 44
FT STRAND 47 53
FT STRAND 58 64
FT STRAND 68 68
FT STRAND 72 88
FT TURN 89 90
FT STRAND 95 97
FT *HELIX 104 110
CORE.
HEADPIECE.
GELSOLIN-LIKE 1.
GELSOLIN-LIKE 2.
GELSOLIN-LIKE 3.
GELSOLIN-LIKE 4.
GELSOLIN-LIKE 5.
GELSOLIN-LIKE 6.
ABSOLUTELY REQUIRED FOR ACTIVITY.
CRUCIAL FOR BINDING AN ACTIN FILAMENT.
POLYPHOSPHOINOSITIDE BINDING (BY
SIMILARITY).
POLYPHOSPHOINOSITIDE BINDING (BY
SIMILARITY).

FT TURN 112 113
FT STRAND 116 118
FT TURN 124 125
SQ SEQUENCE 826 AA; 92479 MW; 6A8898F7DF947389 CRC64;
Query Match 86.1%; Score 31; DB 1; Length 826;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 IFFWI 5
Db 658 IFFWI 662
Search completed: August 6, 2002, 10:40:51
Job time: 248 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 10:40:26 ; Search time 58.49 Seconds
(without alignments)
17.746 Million cell updates/sec

Title: US-09-543-188A-23
Perfect score: 36
Sequence: 1 IFFWIK 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	35	97.2	707	5 Q95XL6	Q95xl6 caenorhabdi
2	34	94.4	158	2 Q93ST6	Q93st6 chlorobium
3	34	94.4	409	2 Q9Z6D5	Q9z6d5 haemophilus
4	33	91.7	62	12 Q91FW6	Q91fw6 chilo iride
5	33	91.7	101	16 Q9AG28	Q9ag28 caulobacter
6	33	91.7	580	2 Q9Z6G8	Q9z6g8 vibrio para
7	32	88.9	73	8 Q36329	Q36329 chondrus cr
8	32	88.9	113	13 Q9PW03	Q9pw03 labeo rohit
9	32	88.9	120	16 Q9CMX2	Q9cmx2 pasteurella
10	32	88.9	174	2 Q9L6G8	Q9l6g8 lactobacill
11	32	88.9	374	5 O16979	O16979 caenorhabdi
12	32	88.9	423	16 Q98QA5	Q98qa5 mycoplasma
13	32	88.9	666	10 Q9FDV6	Q9fdv6 fagus sylv
14	32	88.9	819	12 Q91ET5	Q91ej5 human calic
15	32	88.9	849	12 Q93125	Q93125 human calic
16	32	88.9	988	16 O83345	O83345 treponema p

17	31	86.1	32	9	Q9MCP6	Q9mcp6 bacterioph
18	31	86.1	74	8	Q9MTD6	Q9mtd6 toxoplasma
19	31	86.1	174	2	P97231	P97231 salmonella
20	31	86.1	174	2	P97065	P97065 salmonella
21	31	86.1	174	2	P97066	P97066 salmonella
22	31	86.1	174	2	P97067	P97067 salmonella
23	31	86.1	174	2	P97068	P97068 salmonella
24	31	86.1	180	2	P75018	P75018 escherichia
25	31	86.1	180	2	P71251	P71251 escherichia
26	31	86.1	180	2	P71252	P71252 escherichia
27	31	86.1	180	2	P71253	P71253 escherichia
28	31	86.1	180	2	P71254	P71254 escherichia
29	31	86.1	180	2	P71255	P71255 escherichia
30	31	86.1	180	2	P71256	P71256 escherichia
31	31	86.1	180	2	P71257	P71257 escherichia
32	31	86.1	180	2	P71258	P71258 escherichia
33	31	86.1	180	2	P71259	P71259 escherichia
34	31	86.1	180	2	P71260	P71260 escherichia
35	31	86.1	180	2	P71261	P71261 escherichia
36	31	86.1	190	2	Q9KIA9	Q9kia9 escherichia
37	31	86.1	190	2	Q9KIA8	Q9kia8 escherichia
38	31	86.1	225	9	O64371	O64371 lactobacill
39	31	86.1	258	9	Q9AZP8	Q9azp8 bacterioph
40	31	86.1	262	16	P74256	P74256 synechocyst
41	31	86.1	267	2	Q9A123	Q9a123 candidatus
42	31	86.1	267	16	Q9XIR2	Q9xlr2 thermotoga
43	31	86.1	284	2	Q9AIY2	Q9aiy2 candidatus
44	31	86.1	287	10	O65443	O65443 arabidopsis
45	31	86.1	311	5	Q9T2N0	Q9tzn0 caenorhabdi

ALIGNMENTS

RESULT 1

Q95XL6 ID Q95XL6 PRELIMINARY; PRT; 707 AA.
AC Q95XL6;
DC 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 79.9 KDA PROTEIN.
GN Y73E7A.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Du H., Maupin R.;
RT "The sequence of C. elegans cosmid Y73E7A.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC025727; AAK66029.1; -.
KW Hypothetical protein.
SQ SEQUENCE 707 AA; 79940 MW; 26B924988153DBA8 CRC64;

Query Match 97.2%; Score 35; DB 5; Length 707;

Best Local Similarity 83.3%; Pred. NO. 1.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWIK 6
||||:|
Db 470 IFFWIK 475

RESULT 2

Q93ST6 PRELIMINARY; PRT; 158 AA.
AC Q93ST6
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE BCHF.
GN BCHF.
OS Chlorobium tepidum.
OC Bacteria; Green sulfur bacteria; Chlorobium.
OX NCBI_TaxID=1097;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20433268; PubMed=10976061;
RA Xiong J., Fischer W.M., Inoue K., Nakahara M., Bauer C.E.;
RT "Molecular evidence for the early evolution of photosynthesis."
RL Science 289:1724-1730(2000).
DR EMBL; AY005138; AAG12430.1; -.
SQ SEQUENCE 158 AA; 18396 MW; ABFEF8FCF70B7411 CRC64;

Query Match 94.4%; Score 34; DB 2; Length 158;
Best Local Similarity 83.3%; Pred. NO. 48;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWIK 6
|||||
Db 109 LFFWIK 114

RESULT 3

Q926D5 PRELIMINARY; PRT; 409 AA.
AC Q926D5
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE OAPA.
OS Haemophilus ducreyi.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OX Haemophilus.
OX NCBI_TaxID=730;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=35000;
RX MEDLINE=99115572; PubMed=9916106;
RA Bauer B.A., Lumley S.R., Hansen E.J.;
RT "Characterization of a WaaF (RfaF) homolog expressed by Haemophilus ducreyi."
RL Infect. Immun. 67:899-907(1999).
DR EMBL; AF087414; AAD16054.1; -.
SQ SEQUENCE 409 AA; 45059 MW; 2ACA6861B1ACC451 CRC64;

Query Match 94.4%; Score 34; DB 2; Length 409;
Best Local Similarity 83.3%; Pred. NO. 1.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWIK 6
||||:|
Db 161 IFFWIK 166

RESULT 4

Q91FW6 PRELIMINARY; PRT; 62 AA.
AC Q91FW6
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE 204L.
OS Chilo iridescent virus (CIV) (Insect iridescent virus type 6).
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
OX NCBI_TaxID=10488;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H., Darai G., Fluegel R.M.;
RT "DNA analysis of insect iridescent virus 6: evidence for circular permutation and terminal redundancy."
RL J. Virol. 49:609-614(1984).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86174607; PubMed=3959991;
RA Lorbacher de Ruiz H., Gelderblom H., Hofmann W., Darai G.;
RT "Insect iridescent virus type 6 induced toxic degenerative hepatitis in mice."
RL Med. Microbiol. Immunol. 175:43-53(1986).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=87321126; PubMed=2820141;
RA Schnitzler P., Soltan J.B., Fischer M., Reisner H., Scholz J.,
RA Delius H., Darai G.;
RT "Molecular cloning and physical mapping of the genome of insect iridescent virus type 6: further evidence for circular permutation of the viral genome."
RL Virology 160:66-74(1987).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=89073752; PubMed=3201750;
RA Fischer M., Schnitzler P., Delius H., Darai G.;
RT "Identification and characterization of the repetitive DNA element in the genome of insect iridescent virus type 6."
RL Virology 167:485-496(1988).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=92196996; PubMed=1549908;
RA Handermann M., Schnitzler P., Rosen-Wolff A., Raab K., Sonntag K.C.,
RA Darai G.;
RT "Identification and mapping of origins of DNA replication within the DNA sequences of the genome of insect iridescent virus type 6."
RL Virus Genes 6:19-32(1992).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=93118242; PubMed=1475907;
RA Sonntag K.C., Darai G.;
RT "Characterization of the third origin of DNA replication of the genome of insect iridescent virus type 6."
RL Virus Genes 6:333-342(1992).
RN [7]
RP SEQUENCE FROM N.A.
RX MEDLINE=93260401; PubMed=8492091;
RA Stohwasser R., Raab K., Schnitzler P., Janssen W., Darai G.;
RT "Identification of the gene encoding the major capsid protein of insect iridescent virus type 6 by polymerase chain reaction."
RL J. Gen. Virol. 74:873-879(1993).
RN [8]
RP SEQUENCE FROM N.A.
RX MEDLINE=94167241; PubMed=8121799;
RA Schnitzler P., Hug M., Handermann M., Janssen W., Koonin E.V.,
RA Delius H., Darai G.;
RT "Identification of genes encoding zinc finger proteins, non-histone chromosomal HMG protein homologues, and a putative GTP phosphohydrolase in the genome of Chilo iridescent virus."
RL Nucleic Acids Res. 22:158-166(1994).
RN [9]
RP SEQUENCE FROM N.A.
RX MEDLINE=94353641; PubMed=8073636;

RA Sonntag K.C., Schnitzler P., Koonin E.V., Darai G.;
 RT "Chilo iridescent virus encodes a putative helicase belonging to a
 RT distinct family within the 'DEAD/H' superfamily: implications for the
 RT evolution of large DNA viruses.";
 RL virus Genes 8:151-158(1994).
 RN [10]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95213160; PubMed=7698884;
 RA Sonntag K.C., Schnitzler P., Janssen W., Darai G.;
 RT "Identification of the primary structure and the coding capacity of
 RT the genome of insect iridescent virus type 6 between the genome
 RT coordinates 0.310 and 0.347 (7990 bp).";
 RL Intervirology 37:287-297(1994).
 RN [11]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94292906; PubMed=8021587;
 RA Schnitzler P., Sonntag K.C., Muller M., Janssen W., Bugert J.J.,
 RA Koonin E.V., Darai G.;
 RT "Insect iridescent virus type 6 encodes a polypeptide related to the
 RT largest subunit of eukaryotic RNA polymerase II.";
 RL J. Gen. Virol. 75:1557-1567(1994).
 RN [12]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98141693; PubMed=9482589;
 RA Bahr U., Tidona C.A., Darai G.;
 RT "The DNA sequence of Chilo iridescent virus between the genome
 RT coordinates 0.101 and 0.391: similarities in coding strategy between
 RT insect and vertebrate iridoviruses.";
 RL virus Genes 15:235-245(1997).
 RN [13]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99125223; PubMed=9926400;
 RA Muller K., Tidona C.A., Bahr U., Darai G.;
 RT "Identification of a thymidylate synthase gene within the genome of
 RT Chilo iridescent virus.";
 RL virus Genes 17:243-258(1998).
 RN [14]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99383793; PubMed=10456793;
 RA Muller K., Tidona C.A., Darai G.;
 RT "Identification of a gene cluster within the genome of Chilo
 RT iridescent virus encoding enzymes involved in viral DNA replication
 RT and processing.";
 RL virus Genes 18:243-264(1999).
 RN [15]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21342589; PubMed=11448171;
 RA Jakob N.J., Muller K., Bahr U., Darai G.;
 RT "Analysis of the First Complete DNA Sequence of an Invertebrate
 RT Iridovirus: Coding Strategy of the Genome of Chilo Iridescent Virus.";
 RL Virology 286:182-196(2001).
 RN [16]
 RP SEQUENCE FROM N.A.
 RA Jakob N.J., Mueller K., Bahr U., Darai G.;
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF303741; AAK82056.1; -
 SQ SEQUENCE 62 AA; 7876 MW; 29DF67A85F664B3E CRC64;

Query Match 91.7%; Score 33; DB 12; Length 62;
 Best Local Similarity 83.3%; Pred. No. 29;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IFFWIK 6
 Db 56 IFFWIK 61

RESULT 5
 Q9A628
 ID Q9A628 PRELIMINARY; PRT; 101 AA.
 AC Q9A628;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL PROTEIN CC2267.
 GN CC2267.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 OC Caulobacter.
 OX NCBI_TaxID=69394;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 19089; / CB15;
 RX MEDLINE=21173698; PubMed=11259647;
 RA Niernan W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
 RA Kolony J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
 RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL; AE005897; AAK24238.1; -
 DR TIGR; CC2267; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 101 AA; 11619 MW; 4B5B638942C94AF6 CRC64;

Query Match 91.7%; Score 33; DB 16; Length 101;
 Best Local Similarity 66.7%; Pred. No. 47;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IFFWIK 6
 Db 6 VFFWLK 11

RESULT 6
 Q9Z6G8
 ID Q9Z6G8 PRELIMINARY; PRT; 580 AA.
 AC Q9Z6G8;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE POLAR FLAGELLAR M-RING PROTEIN FLIF.
 GN FLIF.
 OS Vibrio parahaemolyticus.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=670;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=BB22;
 RX MEDLINE=20115548; PubMed=10648530;
 RA Boles B.R., McCarter L.L.;
 RT "Insertional inactivation of genes encoding components of the sodium-
 RT type flagellar motor and switch of vibrio parahaemolyticus.";
 RL J. Bacteriol. 182:1035-1045(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=BB22;
 RA Jacques S., Kim Y.K., McCarter L.L.;
 RT "Components of the polar flagellar switch complex and assembly
 RT apparatus.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF069392; AAD15920.1; -
 DR InterPro; IPR000067; FlgMring_FLIF.
 DR InterPro; IPR002920; YscJ_FLIF.
 DR Pfam; PF01514; YscJ_FlIF; 1.
 DR PRINTS; PR01009; FLGMRINGFLIF.
 KW Flagella.
 SQ SEQUENCE 580 AA; 63813 MW; 4CCE81483A047050 CRC64;

Query Match 91.7%; Score 33; DB 2; Length 580;

Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWIK 6
:|||||
Db 71 IFFWK 76

RESULT 7
Q36329 PRELIMINARY; PRT; 73 AA.
AC Q36329;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE PUTATIVE ORF73 PROTEIN.
GN PUTATIVE ORF73.
OS Chondrus crispus (Carragheen).
OG Mitochondrion.
OC Eukaryota; Rhodophyta; Florideophyceae; Gigartinales; Gigartinaeae;
OC Chondrus.
OX NCBI_TaxID=2769;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=APICES;
RX MEDLINE=95341681; PubMed=7616569;
RA Leblanc C., Boyen C., Richard O., Bonnard G., Grienemberger J.M.,
RA Kloeareg B.;
RT "Complete sequence of the mitochondrial DNA of the rhodophyte Chondrus
RT crispus (Gigartinales). Gene content and genome organization.";
RL J. Mol. Biol. 250:484-495(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=APICES;
RX MEDLINE=95395875; PubMed=7666449;
RA Leblanc C., Kloeareg B., Loiseux-de Goer S., Boyen C.;
RT "DNA sequence, structure and phylogenetic relationship of the
RT mitochondrial small subunit rRNA from the red alga Chondrus crispus
RT (Gigartinales, Rhodophytes).";
RL J. Mol. Evol. 41:196-202(1995).
DR EMBL: 247547; CAA87598.1; -
KW Mitochondrion.
SQ SEQUENCE 73 AA; 8929 MW; 85AC46C006834292 CRC64;

Query Match 88.9%; Score 32; DB 8; Length 73;
Best Local Similarity 83.3%; Pred. No. 52;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IFFWIK 6
:|||||
Db 38 IFFWK 43

RESULT 8
Q9PW03 PRELIMINARY; PRT; 113 AA.
AC Q9PW03;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE ROHU LIVER HYPOTHETICAL PROTEIN (FRAGMENT).
GN LVHP1.
OS Labeo rohita (Indian major carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Labeo.
OX NCBI_TaxID=84645;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Ohri S., Vashishtha A., Ansari Z., Dixit A.;
RT "partial cDNA for a hypothetical protein from Labeo rohita liver.";

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ249232; CAB54039.1; -
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 113
SQ SEQUENCE 113 AA; 13469 MW; 6E36D649DA8FFB08 CRC64;

Query Match 88.9%; Score 32; DB 13; Length 113;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FFWIK 6
:|||||
Db 6 FFWIK 10

RESULT 9
Q9CMX2 PRELIMINARY; PRT; 120 AA.
AC Q9CMX2;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN PM0679.
GN PM0679.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida PM70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL: AE006104; AAK02763.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 120 AA; 14167 MW; 46308C480002FF64 CRC64;

Query Match 88.9%; Score 32; DB 16; Length 120;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FFWIK 6
:|||||
Db 112 FFWIK 116

RESULT 10
Q9L6G8 PRELIMINARY; PRT; 174 AA.
AC Q9L6G8;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE PUTATIVE TRANSPOSASE.
OS Lactobacillus delbrueckii (subsp. bulgaricus).
OG Plasmid pLBB1.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1585;
RN [1]
RP SEQUENCE FROM N.A.
RA Azcarate M.A., Raya R.R.;
RT "Structural analysis of pLBB1, a cryptic plasmid from Lactobacillus
RT delbrueckii subsp. bulgaricus.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF236060; AAF61726.1; -
KW Plasmid.
SQ SEQUENCE 174 AA; 19597 MW; 8DDB65BCDCDB556 CRC64;


```

Query Match      88.9%: Score 32; DB 2; Length 174;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IFFWIK 6
Db 81 LFFWLK 86

RESULT 11
O16979 PRELIMINARY; PRT; 374 AA.
ID O16979
AC O16979
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE HYPOTHETICAL 43.4 KDA PROTEIN.
GN T03D3.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE-99069613; PubMed-9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Murray J., Wohlmann P., Bauer C., Biewald T.;
RT "The sequence of C. elegans cosmid T03D3."
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF022980; AAC24188.1;
KW HYPOTHETICAL protein.
SQ SEQUENCE 374 AA; 43420 MW; DEEE25C5B9E05B5F CRC64;

Query Match      88.9%: Score 32; DB 5; Length 374;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FFWIK 6
Db 36 FFWIK 40

RESULT 12
Q98QA5 PRELIMINARY; PRT; 423 AA.
ID Q98QA5
AC Q98QA5;
DT 01-OCT-2001 (TRENBLrel. 18, Created)
DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TRENBLrel. 18, Last annotation update)
DE HYPOTHETICAL PROTEIN MYPU_4610.
GN MYPU_4610.
OS Mycoplasma pulmonis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2107;
RN [1]
RP SEQUENCE FROM N.A.

Query Match      88.9%: Score 32; DB 10; Length 666;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FFWIK 6
Db 390 LFFWLK 395

RESULT 13
Q9FDV6 PRELIMINARY; PRT; 666 AA.
ID Q9FDV6
AC Q9FDV6
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE PROTEIN KINASE.
GN PK2.
OS Fagus sylvatica (Beechnut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
OC euroids 1; Fagales; Fagaceae; Fagus.
OX NCBI_TaxID=28930;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SEED;
RA Lorenzo O., Rodriguez D., Nicolas C., Nicolas G.;
RT "Characterization and expression of two protein kinase genes and EIN-3
like gene, which are regulated by ABA and GA3 in dormant Fagus
sylvatica seeds."
RL (In) Black M., Bradford K.J., Vazquez-Ramos J. (eds.);
RL SED BIOLOGY-ADVANCES AND APPLICATIONS 32, pp.329-340,
RL CAB International, Oxfordshire, UK (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=SEED;
RA Lorenzo O.;
RT Thesis (2000), Fisiologia Vegetal, Facultad de Biologia,
Universidade de Salamanca.
CC -i- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AJ298993; CAC09581.1;
DR InterPro; IPR000719; Euk.pkinase.
DR InterPro; IPR002290; Ser.thr.pkinase.
DR InterPro; IPR001245; Tyr.pkinase.
DR Pfam; PF00069; pkinase; 2.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00108; Serine/threonine-protein kinase; Transferase.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 666 AA; 74385 MW; DA58681276259C97 CRC64;

Query Match      88.9%: Score 32; DB 16; Length 423;
Best Local Similarity 66.7%; Pred. No. 2.9e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IFFWIK 6
Db 390 LFFWLK 395

RESULT 13
Q9FDV6 PRELIMINARY; PRT; 666 AA.
ID Q9FDV6
AC Q9FDV6
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE PROTEIN KINASE.
GN PK2.
OS Fagus sylvatica (Beechnut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
OC euroids 1; Fagales; Fagaceae; Fagus.
OX NCBI_TaxID=28930;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SEED;
RA Lorenzo O., Rodriguez D., Nicolas C., Nicolas G.;
RT "Characterization and expression of two protein kinase genes and EIN-3
like gene, which are regulated by ABA and GA3 in dormant Fagus
sylvatica seeds."
RL (In) Black M., Bradford K.J., Vazquez-Ramos J. (eds.);
RL SED BIOLOGY-ADVANCES AND APPLICATIONS 32, pp.329-340,
RL CAB International, Oxfordshire, UK (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=SEED;
RA Lorenzo O.;
RT Thesis (2000), Fisiologia Vegetal, Facultad de Biologia,
Universidade de Salamanca.
CC -i- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AJ298993; CAC09581.1;
DR InterPro; IPR000719; Euk.pkinase.
DR InterPro; IPR002290; Ser.thr.pkinase.
DR InterPro; IPR001245; Tyr.pkinase.
DR Pfam; PF00069; pkinase; 2.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00108; Serine/threonine-protein kinase; Transferase.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 666 AA; 74385 MW; DA58681276259C97 CRC64;

```

Db 369 FFWIK 373

RESULT 14

Q91EJ5 PRELIMINARY; PRT; 819 AA.
AC Q91EJ5; (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE POLYPROTEIN (FRAGMENT).
OS Human calicivirus Hu/SLV/Lyon/598/97/F.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Sapporo-like viruses.
OX NCBI_TaxID=144196;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/SLV/LYON/598/97/F;
RA Schuffenecker I., Ando T., Thouvenot D., Floret D.;
RT "Molecular characterization of Sapporo-like viruses (SLV) in fecal
RT samples from gastroenteritis cases in Lyon, France."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ271056; CAC41375.1;
FT NON_TERMINAL
SQ SEQUENCE 819 AA; 88180 MW; 7B8055772C8119CA CRC64;

Query Match 88.9%; Score 32; DB 12; Length 819;

Best Local Similarity 100.0%; Pred. No. 5.5e+02; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FFWIK 6

Db 163 FFWIK 167

RESULT 15

O93125 PRELIMINARY; PRT; 849 AA.
AC O93125;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE POLYPROTEIN (FRAGMENT).
OS Human calicivirus isolates.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=150080;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SAPPORO-LIKE;
RX MEDLINE=98336510; PubMed=9672639;
RA Jiang X., Cubitt W.D., Berke T., Zhong W., Dai X., Nakata S.,
RA Pickering L.K., Matson D.O.;
RT "Sapporo-like human caliciviruses are genetically and antigenically
RT diverse."
RL Arch. Virol. 142:1813-1827(1997).
DR EMBL; U95645; AAC40584.1;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR004004; Calici_pol_hel.
DR Pfam; PF00915; Calici_coat_2;
DR PRINTS; PR00318; CALICIVIRUSNS.
FT NON_TERMINAL
FT CHAIN 1 292 RNA DEPENDENT RNA POLYMERASE.
FT CHAIN 293 849 CAPSID PROTEIN.
SQ SEQUENCE 849 AA; 91906 MW; 941C52A621EA70 CRC64;

Query Match 88.9%; Score 32; DB 12; Length 849;

Best Local Similarity 100.0%; Pred. No. 5.7e+02; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FFWIK 6

Db 196 FFWIK 200

Search completed: August 6, 2002, 10:40:28
Job time: 225 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 10:38:48 ; Search time 73.38 Seconds
(without alignments)
25.733 Million cell updates/sec

Title: us-09-543-188a-34

Perfect score: 116

Sequence: 1 DGHPOGWGGHPQGWG 17

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database

A_Geneseq_032802.*

- 1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	58	50.0	258	22	AAB72370 Marsupial prion pr
2	56.5	48.7	179	21	AAB42770 Human ORFX ORF2534
3	56.5	48.7	453	22	AAM25915 Human protein sequ
4	56.5	48.7	786	20	AAV55937 Human SULU3 protei
5	56.5	48.7	1001	20	AAV55942 Human/Murine SULU3
6	56.5	48.7	1001	20	AAV49896 Rat TAO1 kinase
7	55	47.4	524	22	AAV40332 Human polypeptide
8	55	47.4	593	21	AAB43002 Human ORFX ORF2766
9	55	47.4	593	22	AAM40333 Human polypeptide
10	54	46.6	159	22	ABG63348 Drosophila melanog
11	54	46.6	181	22	ABG69088 Drosophila melanog

12	54	46.6	334	21	AAG36181 Arabidopsis thalia
13	54	46.6	375	21	AAG36180 Arabidopsis thalia
14	54	46.6	441	21	AAG36179 Arabidopsis thalia
15	53.5	46.1	236	22	AAE00342 Rat secreted facto
16	53.5	46.1	292	22	AAE00343 Rat secreted facto
17	53	45.7	77	22	AAU22202 Human cardiopascu
18	53	45.7	137	22	ABG69236 Drosophila melanog
19	53	45.7	228	22	AAE05350 Mouse surface glyc
20	53	45.7	325	22	ABG71023 Drosophila melanog
21	53	45.7	477	21	AAB08774 Amino acid sequenc
22	52.5	45.3	107	21	AAB40288 Human ORFX ORF52 p
23	52.5	45.3	258	21	AAG51723 Arabidopsis thalia
24	52	44.8	274	21	AAG54075 Arabidopsis thalia
25	52	44.8	304	21	AAG54075 Arabidopsis thalia
26	52	44.8	322	21	AAG54074 Arabidopsis thalia
27	52	44.8	415	21	AAG14062 Arabidopsis thalia
28	52	44.8	445	21	AAG14061 Arabidopsis thalia
29	52	44.8	463	21	AAG14060 Arabidopsis thalia
30	51.5	44.4	178	22	AAU34855 E. coli cellular pr
31	51.5	44.4	285	21	AAU34855 Human secreted pro
32	51.5	44.4	535	22	AAE10203 Human bone marrow
33	51.5	44.4	1684	22	ABG12028 Human dJ1042K10.4
34	51.5	44.4	1684	22	AAU79579 Human protein SEQ
35	51.5	44.4	1723	22	AAU78595 Human protein SEQ
36	51	44.0	118	22	AAO12123 Human polypeptide
37	51	44.0	148	22	AAO10299 Human polypeptide
38	51	44.0	166	22	AAU29635 Novel human secret
39	51	44.0	325	22	AAU29142 Human PRO polypept
40	51	44.0	325	22	AAU78557 Human PRO3566. Ho
41	51	44.0	373	20	AAU34998 Chlamydia pneumoni
42	51	44.0	557	22	AAU30324 Human cytoskeleton
43	50	43.1	26	14	AAU38035 Prion protein regi
44	50	43.1	70	21	AAG13426 Arabidopsis thalia
45	50	43.1	75	22	AAU52734 Propionibacterium

ALIGNMENTS

RESULT 1

AAB72370

ID AAB72370 standard; Peptide: 258 AA.

AC AAB72370;

DT 17-MAY-2001 (first entry)

XX Marsupial prion protein cellular form (PrPc) amino acid sequence.

DE Prion protein; cellular form; PrPc; stable region; antibody; BSE; CJD;

XX prion disease; spongiform encephalopathies; Scrapie; marsupial;

KW bovine spongiform encephalopathy; BSE; Creutzfeldt-Jakob disease.

XX Marsupialia.

OS Marsupialia.

XX Key Location/Qualifiers

XX Region 176..221

XX FT /note= "Stable region, specifically claimed in claim 3"

XX WO200107479-A2.

XX PD 01-FEB-2001.

XX PF 25-JUL-2000; 2000WO-GB02873.

XX PR 27-JUL-1999; 99GB-0017491.

XX PR 30-JUL-1999; 99GB-0017878.

XX PA (INCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

XX PI Collinge J, Clarke AR, Waltho JP, Jackson GS, Hosszu LLP;

XX DR WPI; 2001-16853E/17.

XX PD 26-JUL-2001.
 XX PF 22-DEC-2000; 2000WO-US35017.
 XX PR 23-DEC-1999; 99US-0471275.
 XX PR 21-JAN-2000; 2000US-0488725.
 XX PR 25-APR-2000; 2000US-0552317.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Tang YT, Liu C, Drmanac RT;
 XX WPI; 2001-457603/49.
 XX DR N-PSDB; AAH99856.
 XX PT Isolated human polynucleotides encoding polypeptides, useful for the
 PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
 XX PS Claim 20; Page 290; 1217pp; English.
 XX CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
 CC AAM25963. The proteins can have activities based on the tissues and
 CC cells they are expressed in, such as: antinflammatory; antirheumatic;
 CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
 CC central nervous system; viricide; anti-HIV; fungicide; antimutagen;
 CC cardiovascular; antianaemic; antiagregant; haemostatic; vulnerary;
 CC antilucer; osteopathic; dermatological; antiallergic; antiasthmatic;
 CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antisense therapy and vaccine
 CC production. The proteins and polynucleotides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis
 CC of disorders associated with the activity of a protein e.g. inflammation,
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders.
 XX SQ Sequence 453 AA;
 Query Match 48.78; Score 56.5; DB 22; Length 453;
 Best Local Similarity 57.98; Pred. No. 9.6;
 Matches 11; Conservative 0; Mismatches 5; Indels 3; Gaps 1;
 Qy 2 GGHPPQGWG---GGHPQGWG 17
 || || || | || ||
 Db 417 ggpqagwhpmagggppqpwg 435
 RESULT 4
 AAY55937
 ID AAY55937 standard; Protein; 786 AA.
 XX AC AAY55937;
 XX 18-FEB-2000 (first entry)
 XX Human SULU3 protein.
 XX Antirheumatic; antiarthritic; antinflammatory; antiallergic; osteopathic;
 XX antipsoriatic; antiarteriosclerotic; antiasthmatic; immunosuppressive;
 XX neuroprotective; cardiant; cerebroprotective; cytostatic; antidiabetic;
 XX vulnery; STE20; protein kinase; STLK2; STLK3; STLK4; STLK5; STLK6; STLK7;
 XX ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4; PAK5; antagonist;
 XX antibody; gene therapy; rheumatoid arthritis; artherosclerosis; asthma;
 XX inflammatory bowel disease; Crohn's disease; osteoarthritis; psoriasis;
 XX rhinitis; autoimmunity; organ transplantation; multiple sclerosis;

KW myocardial infarction; cardiovascular disease; stroke; renal failure;
 KW oxidative stress-related neurodegenerative disorder; Parkinson's disease;
 KW amyotrophic lateral sclerosis; Leigh syndrome; cancer; cardiomyopathy;
 KW ischemic disorder; inflammation; diabetes mellitus; fibrosis; mitosis;
 KW mesangial disorder; growth regulation; wound healing; T cell activation;
 KW immunosuppressant.
 XX OS Homo sapiens.
 XX PN WO9953036-A2.
 XX PD 21-OCT-1999.
 XX 13-APR-1999; 99WO-US08150.
 XX PF 14-APR-1998; 98US-0081784.
 XX PR (SUGE-) SUGEN INC.
 XX PA Plowman G, Martinez R, Whyte D;
 XX PI WPI; 1999-611301/52.
 XX DR N-PSDB; AAZ40489.
 XX Novel kinase-related polypeptides used for the diagnosis and treatment
 XX of kinase-related diseases and disorders -
 XX Claim 11; Page 296-299; 387pp; English.
 XX This sequence represents a novel STE20-related protein kinase. The
 CC invention relates to nucleic acid molecule encoding a kinase polypeptide
 CC selected from STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3,
 CC ZC4, KHS2, SULU1, SULU3, GEK2, PAK4 and PAK5. The proteins are used to
 CC identify agonists and antagonists, and to raise antibodies. The
 CC polynucleotides are useful in gene therapy protocols. The polynucleotides,
 CC polypeptides, antibodies, antagonists and agonists may be used to treat
 CC diseases such as immune-related disorders and diseases (e.g. rheumatoid
 CC arthritis, artherosclerosis, chronic inflammatory bowel disease (e.g.
 CC Crohn's disease), asthma, osteoarthritis, psoriasis, atherosclerosis,
 CC rhinitis, autoimmunity, and organ transplantation, chronic inflammatory
 CC pelvic disease, multiple sclerosis, organ transplantation, myocardial
 CC infarction, cardiovascular disease, stroke, renal failure, oxidative
 CC stress-related neurodegenerative disorders (e.g. amyotrophic lateral
 CC sclerosis, Parkinson's disease and Leigh syndrome), cancer,
 CC cardiomyopathies, ischemic disorders, inflammatory disorders, diabetes
 CC mellitus, fibrotic and mesangial disorders. The proteins may also be
 CC useful for cell-growth regulation (e.g. in wound healing), T cell
 CC activation, mitosis control, and as immunosuppressants.
 XX SQ Sequence 786 AA;
 Query Match 48.78; Score 56.5; DB 20; Length 786;
 Best Local Similarity 57.98; Pred. No. 16;
 Matches 11; Conservative 0; Mismatches 5; Indels 3; Gaps 1;
 Qy 2 GGHPPQGWG---GGHPQGWG 17
 || || || | || ||
 Db 712 ggpqagwhpmagggppqpwg 730
 RESULT 5
 AAY55942
 ID AAY55942 standard; Peptide; 1001 AA.
 XX AC AAY55942;
 XX 18-FEB-2000 (first entry)
 XX Human/Murine SULU3 consensus protein sequence.
 XX Antirheumatic; antiarthritic; antinflammatory; antiallergic; osteopathic;
 KW antipsoriatic; antiarteriosclerotic; antiasthmatic; immunosuppressive;

KW neuroprotective; cardiant; cerebroprotective; cytotatic; antidiabetic;
 KW vulnery; STE20; protein kinase; STLK2; STLK3; STLK4; STLK5, STLK6, STLK7;
 KW ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4; PAK5; antagonist;
 KW antibody; gene therapy; rheumatoid arthritis; artherosclerosis; asthma;
 KW inflammatory bowel disease; Crohn's disease; osteoarthritis; psoriasis;
 KW rhinitis; autoimmunity; organ transplantation; multiple sclerosis;
 KW myocardial infarction; cardiovascular disease; stroke; renal failure;
 KW oxidative stress-related neurodegenerative disorder; Parkinson's disease;
 KW amyotrophic lateral sclerosis; Leigh syndrome; cancer; cardiomyopathy;
 KW ischemic disorder; inflammation; diabetes mellitus; fibrosis; mitosis;
 KW mesangial disorder; growth regulation; wound healing; T cell activation;
 KW immunosuppressant.

XX Homo sapiens.

OS Mus sp.

XX WO9953036-A2.

PN 21-OCT-1999.

XX 13-APR-1999; 99WO-US08150.

PF 14-APR-1998; 98US-0081784.

PR (SUGEN-) SUGEN INC.

XX Plowman G, Martinez R, Whyte D;

PI WPI; 1999-611301/52.

XX Novel kinase-related polypeptides used for the diagnosis and treatment

PT of kinase-related diseases and disorders

XX Claim 11; Page 312-315; 387pp; English.

XX This sequence represents a consensus peptide sequence contained in novel

CC STK20-related protein kinases. The invention relates to a nucleic acid

CC molecule encoding a kinase polypeptide selected from STLK2, STLK3, STLK4,

CC STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4

CC and PAK5. The proteins are used to identify agonists and antagonists, and

CC to raise antibodies. The polynucleotides are useful in gene therapy

CC protocols. The polynucleotides, polypeptides, antibodies, antagonists and

CC agonists may be used to treat diseases such as immune-related disorders

CC and diseases (e.g. rheumatoid arthritis, artherosclerosis, chronic

CC inflammatory bowel disease (e.g. Crohn's disease), asthma,

CC osteoarthritis, psoriasis, atherosclerosis, rhinitis, autoimmunity,

CC and organ transplantation, chronic inflammatory pelvic disease, multiple

CC sclerosis, organ transplantation, myocardial infarction, cardiovascular

CC disease, stroke, renal failure, oxidative stress-related

CC neurodegenerative disorders (e.g. amyotrophic lateral sclerosis,

CC Parkinson's disease and Leigh syndrome), cancer, cardiomyopathies,

CC ischemic disorders, inflammatory disorders, diabetes mellitus, fibrotic

CC and mesangial disorders. The proteins may also be useful for cell growth

CC regulation (e.g. in wound healing), T cell activation, mitosis control,

CC and as immunosuppressants.

XX Sequence 1001 AA;

SQ

Query Match 48.7%; Score 56.5; DB 20; Length 1001;
 Best Local Similarity 57.9%; Pred. No. 21;
 Matches 11; Conservative 0; Mismatches 5; Indels 3; Gaps 1;

OY 2 GGHPQGWG---GGHPQGWG 17

||| ||| ||| ||| |||

Db 927 g9ppqawghpmgpggpgw 945

RESULT 6

AA49896

ID AA49896 standard; Protein: 1001 AA.

XX

AC AA49896;

XX 27-JAN-2000 (first entry)
 DT Rat TAO1 kinase.
 DE

XX TAO1; TAO2; MEK3; mitogen activated protein kinase; phosphorylation;
 KW p38; protein kinase; cancer; inflammation; autoimmune disease;
 KW degeneration; insulin-resistant diabetes; metabolic disorder;
 KW neurodegeneration; MAP kinase; MAP/ERK kinase.

XX Rattus sp.

XX WO9953076-A1.

PN 21-OCT-1999.

XX 14-APR-1999; 99WO-US08165.

PF 14-APR-1998; 98US-0060410.

PR (TEXA) UNIV TEXAS SYSTEM.

XX Cobb M, Hutchison M, Chen Z, Berman K;

PI WPI; 1999-633831/54.

XX N-PSDB; AA232435.

XX New polypeptides that phosphorylate kinase, used to screen for

PT modulators for treating e.g. cancer or inflammation

XX Claim 1; Fig 1; 95pp; English.

XX The present sequence represents rat TAO1 protein kinase, which is capable

CC of phosphorylating MEK3 (a MAP/ERK kinase). TAO kinases, and related

CC polypeptides, are used to screen for modulators of stress-responsive

CC mitogen activated protein (MAP) kinase pathways. These modulators are

CC potentially useful for treating or preventing (1) inflammation,

CC autoimmune disease, cancer and degeneration (inhibitors of

CC phosphorylation); or (2) insulin-resistant diabetes, metabolic disorders

CC and neurodegeneration (enhancers of phosphorylation). TAO kinases are

CC also used to raise specific antibodies, useful therapeutically as

CC modulators and as immunoassay reagents for detecting TAO kinases.

CC TAO kinase polynucleotides can be used: (a) for recombinant expression

CC of TAO kinases; and (b) in the form of fragments, for detecting TAO

CC kinase polynucleotides in standard hybridisation and amplification

CC tests. TAO kinases are highly specific for MEK3.

XX Sequence 1001 AA;

SQ

Query Match 48.7%; Score 56.5; DB 20; Length 1001;
 Best Local Similarity 57.9%; Pred. No. 21;
 Matches 11; Conservative 0; Mismatches 5; Indels 3; Gaps 1;

OY 2 GGHPQGWG---GGHPQGWG 17

||| ||| ||| ||| |||

Db 927 g9ppqawghpmgpggpgw 945

RESULT 7

AA40332

ID AA40332 standard; Protein: 524 AA.

XX

AC AA40332;

XX 22-OCT-2001 (first entry)

DT Human polypeptide SEQ ID NO 3477.

DE

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;

KW peripheral nervous system; neuropathy; central nervous system; CNS;

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW	leukaemia.
XX	
XX	Homo sapiens.
XX	
PN	WO200153312-A1.
XX	
PD	26-JUL-2001.
XX	
PF	26-DEC-2000; 200OWO-US34263.
XX	
PR	21-JAN-2000; 2000US-0488725.
PR	25-APR-2000; 2000US-0552317.
PR	09-JUL-2000; 2000US-0598042.
PR	19-JUL-2000; 2000US-0620312.
PR	03-AUG-2000; 2000US-0653450.
PR	14-SEP-2000; 2000US-0662191.
PR	19-OCT-2000; 2000US-0693036.
PR	29-NOV-2000; 2000US-0727344.
PA	(HYSE-) HYSEQ INC.
XX	
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI	Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX	
DR	WPI: 2001-442253/47.
XX	
DR	N-PSDB; AAI59488.
XX	
PT	Novel nucleic acids and polypeptides, useful for treating disorders
PT	such as central nervous system injuries .
XX	
PS	Example 6; SEQ ID NO 3477; 10078pp; English.
XX	
CC	The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC	the encoded polypeptides (AAK38642-AAK42213) with nootropic,
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful
CC	in gene therapy. A composition containing a polypeptide or polynucleotide
CC	of the invention may be used to treat diseases of the peripheral nervous
CC	system, such as peripheral nervous injuries, peripheral neuropathy and
CC	localised neuropathies and central nervous system diseases, such as
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC	utilisation of the activities such as: Immune system suppression,
CC	activation/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC	assays for receptor activity, arthritis and inflammation, leukaemias and
CC	C.N.S disorders.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification.
XX	
SQ	Sequence 524 AA;
	Query Match 47.4%; Score 55; DB 22; Length 524;
	Best Local Similarity 69.2%; Pred. No. 17;
	Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0
QY	2 GGHPOGWGGGHPQ 14
Db	56 gqpppgwggghpq 68
RESULT	8
ID	AAB43002
AC	ID AAB43002 standard; Protein; 593 AA.
XX	
AC	AAB43002;
XX	
DT	08-FEB-2001 (first entry)
XX	
DE	Human ORFX ORE2766 polypeptide sequence SEQ ID NO:5532.
XX	

PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.

PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161952.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 46.6%; Score 54; DB 21; Length 334;

Best Local Similarity 62.5%; Pred. No. 15;

Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 2 GGHPQGGGGHPQGGW 17
Db 225 gfgpgmggmpagmg 240

RESULT 13

AAG36180

ID AAG36180 standard; Protein; 375 AA.

XX AAG36180;

AC AAG36180;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 44303.

XX Protein identification; signal transduction pathway; metabolic pathway;

XX hybridisation assay; genetic mapping; gene expression control; promoter;

XX termination sequence.

XX Arabidopsis thaliana.

OS EPI033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140334.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142134.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.

```
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160983.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 46.6%; Score 54; DB 21; Length 375;
Best Local Similarity 62.5%; Pred. No. 17;
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GCHPOGCGGGHPQGWG 17
Db 266 gfpbgmggmpagmg 281

RESULT 14
AAG36179
ID AAG36179 standard; Protein; 441 AA.
XX AC AAG36179;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID No: 44302.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
```

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151086.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 46.6%; Score 54; DB 21; Length 441;
Best Local Similarity 62.5%; Pred. No. 20;
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 GGHPQGGGGHPQGGW 17
||| | ||| | | |
Db 332 gffpggggmgmgmg 347

RESULT 15
AAE00342
ID AAE00342 standard; Protein; 236 AA.
XX AC AAE00342;
XX DT 19-JUN-2001 (first entry)
XX DE Rat secreted factor protein #1 encoded by clone P00188_p12.
XX KW Rat; secreted factor; clone P00188_p12; cardiac; antiinflammatory;
KW antiarrhythmic; antiarteriosclerotic; antiatherosclerotic; nephropathic;
KW antidiabetic; immunosuppressive; antiasthmatic; antirheumatoid;
KW antibacterial; osteopathic; cerebroprotective; vasotropic; antileuc;
KW neurotropic; neuroprotective; congestive heart failure; myocarditis;
KW hypertrophic cardiomyopathy; angina pectoris; myocardial infarction;
KW kidney disease; acute renal failure; renal glucosuria; renal infarction;
KW polycystic kidney disease; hereditary nephritis; inflammatory disease;
KW tumour angiogenesis; osteoarthritis; toxic shock syndrome; psoriasis;
KW stroke; neural trauma; cerebral malaria; Crohn's disease; osteoporosis;
KW ulcerative colitis; Alzheimer's disease; gene therapy.
XX OS Rattus norvegicus.
XX FH Key Location/Qualifiers
FT Peptide 1..24
FT /label= Signal_peptide
FT Protein 25..236
FT /label= Rat_mature_secreted_factor_protein
FT Region 68..74
FT /label= Ly-6/u-PAR_domain_signature
FT Region 136..144
FT /label= Ly-6/u-PAR_domain_signature
FT Domain 215..235
FT /label= Transmembrane_domain
XX WO200123564-A1.
XX 05-APR-2001.

XX 27-SEP-2000; 2000WO-US26544.
XX
XX 27-SEP-1999; 99US-0156280.
XX
XX (SCIO-) SCIOS INC.
XX
XX Stanton LW, Kapoun AM;
XX
XX WPI; 2001-266159/27.
XX N-PSDB; AAD03557.
XX
XX Novel secreted factor encoded by clone P00188D12 which is
XX differentially expressed in certain disease states, useful in
XX diagnosing and treating cardiac, renal or inflammatory diseases
XX
XX Claim 1; Fig 1; 7lpp; English.
XX
XX The patent discloses novel secreted factor protein encoded by clone
XX P00188_D12. The secreted factor is differentially expressed in certain
XX disease states. Secreted protein, its antibodies, antagonists or
XX compositions comprising them are useful in the diagnosis and treatment
XX of cardiac diseases such as congestive heart failure, myocarditis,
XX hypertrophic cardiomyopathy, angina pectoris, myocardial infarction,
XX cardiac arrhythmia, arteriosclerosis, kidney diseases such as acute
XX renal failure, renal glucosuria; renal infarction, nephrogenic
XX diabetes insipidus, polycystic kidney disease, hereditary nephritis
XX and inflammatory diseases such as asthma, autoimmune diabetes, tumour
XX angiogenesis, rheumatoid arthritis, osteoarthritis, toxic shock
XX syndrome, asthma, stroke, neural trauma, psoriasis, cerebral malaria,
XX osteoporosis, Crohn's disease, ulcerative colitis, Alzheimer's disease.
XX Secreted protein DNA is useful in antisense-mediated gene inhibition
XX and in gene therapy. An array comprising one or more oligonucleotides
XX complementary to reference RNA or DNA encoding the secreted factor is
XX useful for detecting cardiac, kidney and inflammatory disease.
XX The present sequence is rat secreted factor protein encoded by
XX P00188_D12 clone.
XX
XX Sequence 236 AA;
SQ

Query Match 46.1%; Score 53.5; DB 22; Length 236;
Best Local Similarity 44.0%; Pred. No. 12;
Matches 11; Conservative 0; Mismatches 5; Indels 9; Gaps 1;
Qy 1 DGGH-----PQGWGGHPQGW 16
||| ||| ||| ||| |||
Db 191 dqahlskqgkanpqngagpagw 215

Search completed: August 6, 2002, 10:38:49
Job time: 126 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 10:39:22 ; Search time 26.75 seconds
(without alignments)
15.523 Million cell updates/sec

Title: US-09-543-188A-34

Perfect score: 116

Sequence: 1 DGGHPQGWGGHPQGWG 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 segs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56.5	48.7	1001	4	US-09-060-410-2
2	53	45.7	477	3	US-09-118-324-2
3	50	43.1	26	1	US-08-244-701B-46
4	50	43.1	208	4	US-09-128-450-18
5	50	43.1	208	4	US-09-823-494-18
6	50	43.1	253	1	US-08-242-188-2
7	50	43.1	253	1	US-08-509-261A-2
8	50	43.1	253	1	US-08-660-626-8
9	50	43.1	253	1	US-08-692-892-2
10	50	43.1	253	2	US-08-713-939A-2
11	50	43.1	253	2	US-08-868-162A-22
12	50	43.1	253	4	US-09-031-168-8
13	50	43.1	253	4	US-09-128-450-20
14	50	43.1	253	4	US-09-036-579-2
15	50	43.1	253	4	US-09-823-494-20
16	50	43.1	254	4	US-09-128-450-26
17	50	43.1	254	4	US-09-823-494-26
18	50	43.1	255	1	US-08-242-188-4
19	50	43.1	255	1	US-08-509-261A-4
20	50	43.1	255	1	US-08-660-626-10
21	50	43.1	255	1	US-08-692-892-4
22	50	43.1	255	2	US-08-713-939A-4
23	50	43.1	255	2	US-08-868-162A-24
24	50	43.1	255	4	US-09-031-168-10
25	50	43.1	255	4	US-09-036-579-4
26	50	43.1	256	4	US-09-128-450-22
27	50	43.1	256	4	US-09-823-494-22

28	50	43.1	263	1	US-08-242-188-3	Sequence 3, Appli
29	50	43.1	263	1	US-08-509-261A-3	Sequence 3, Appli
30	50	43.1	263	1	US-08-660-626-9	Sequence 9, Appli
31	50	43.1	263	1	US-08-692-892-3	Sequence 3, Appli
32	50	43.1	263	2	US-08-713-939A-3	Sequence 3, Appli
33	50	43.1	263	2	US-08-868-162A-23	Sequence 23, Appli
34	50	43.1	263	4	US-09-031-168-9	Sequence 9, Appli
35	50	43.1	263	4	US-09-036-579-3	Sequence 3, Appli
36	50	43.1	264	4	US-09-128-450-21	Sequence 21, Appli
37	50	43.1	264	4	US-09-823-494-21	Sequence 21, Appli
38	50	43.1	420	2	US-08-845-998-8	Sequence 8, Appli
39	50	43.1	420	4	US-09-206-537-8	Sequence 8, Appli
40	50	43.1	420	4	US-09-430-854-8	Sequence 8, Appli
41	49	42.2	27	2	US-08-859-201-12	Sequence 12, Appli
42	48.5	41.8	254	1	US-08-242-188-1	Sequence 1, Appli
43	48.5	41.8	254	1	US-08-509-261A-1	Sequence 1, Appli
44	48.5	41.8	254	1	US-08-660-626-7	Sequence 7, Appli
45	48.5	41.8	254	1	US-08-692-892-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-060-410-2
; Sequence 2, Application US/09060410
; Patent No. 6165461
; GENERAL INFORMATION:
; APPLICANT: Cobb, Melanie
; APPLICANT: Hutchinson, Michele
; APPLICANT: Chen, Zhu
; APPLICANT: Berman, Kevin
; TITLE OF INVENTION: TAO PROTEIN KINASES AND METHODS OF USE
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/060,410
; APPLICATION NUMBER: US/09/060,410
; FILING DATE: 14-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 860098.421
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1001 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-060-410-2

Query Match 48.7%; Score 56.5; DB 4; Length 1001;
Best Local Similarity 57.9%; Pred. No. 8.7;
Matches 11; Conservative 0; Mismatches 5; Indels 3; Gaps 1;
QY 2 GGHPQGWG---GGHPQGWG 17
||| ||||| ||| |||

Db 927 GCTPOAGHPMOGGPQWG 945

RESULT 2

US-09-118-324-2
; Sequence 2, Application US/09118324
; Patent No. 6110720
; GENERAL INFORMATION:
; APPLICANT: Li, Xin-Liang
; APPLICANT: Ljungdahl, Lars G.
; APPLICANT: Chen, Huizhong
; TITLE OF INVENTION: Orpinomyces Cellulase CelE Protein and Coding Sequences
; FILE REFERENCE: 32-98sequence listing
; CURRENT APPLICATION NUMBER: US/09/118.324
; CURRENT FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Orpinomyces sp. PC-2
US-09-118-324-2

Query Match 45.7%; Score 53; DB 3; Length 477;
Best Local Similarity 53.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 HPOGWWGGHPQGW 16

||| ||| :||

Db 180 HPNEWGGDQGW 192

RESULT 3

US-08-244-701B-46
; Sequence 46, Application US/08244701B
; Patent No. 5773572
; GENERAL INFORMATION:
; APPLICANT: Fishleigh, Robert V.
; APPLICANT: Robson, Barry
; APPLICANT: Mee, Roger P.
; TITLE OF INVENTION: Fragments of Prion Proteins
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/244.701B
; FILING DATE: 02-JUN-1994
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Fanucci, Allan A.
; REGISTRATION NUMBER: 30,256
; REFERENCE/DOCKET NUMBER: 8080-007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: peptide
US-08-244-701B-46

Query Match 43.1%; Score 50; DB 1; Length 26;
Best Local Similarity 57.1%; Pred. No. 1.5;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 HPOGWWGGHPQGW 17

||| ||| :||

Db 4 HGGWGGPHGGGW 17

RESULT 4

US-09-128-450-18
; Sequence 18, Application US/09128450
; Patent No. 6211149
; GENERAL INFORMATION:
; APPLICANT: Chesebro, Bruce W
; APPLICANT: Caughey, Byron W
; APPLICANT: Chabry, Joelle
; APPLICANT: Priola, Susette
; TITLE OF INVENTION: Inhibitors of Formation of Protease Resistant Prion
; TITLE OF INVENTION: Protein
; FILE REFERENCE: 50121
; CURRENT APPLICATION NUMBER: US/09/128.450
; CURRENT FILING DATE: 1998-08-03
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Hamster sp.
; FEATURE:
; NAME/KEY: NON_TER
; LOCATION: (1)..(2)
US-09-128-450-18

Query Match 43.1%; Score 50; DB 4; Length 208;
Best Local Similarity 57.1%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 HPOGWWGGHPQGW 17

||| ||| :||

Db 39 HGGWGGPHGGGW 52

RESULT 5

US-09-823-494-18
; Sequence 18, Application US/09823494
; Patent No. 6355610
; GENERAL INFORMATION:
; APPLICANT: Chesebro, Bruce W
; APPLICANT: Caughey, Byron W
; APPLICANT: Chabry, Joelle
; APPLICANT: Priola, Susette
; TITLE OF INVENTION: Inhibitors of Formation of Protease Resistant Prion
; TITLE OF INVENTION: Protein
; FILE REFERENCE: 50121
; CURRENT APPLICATION NUMBER: US/09/823.494
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/128,450
; PRIOR FILING DATE: 1998-08-03
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Hamster sp.
; FEATURE:
; NAME/KEY: NON_TER
; LOCATION: (1)..(2)

US-09-823-494-18

Query Match 43.1% Score 50; DB 4; Length 208;
Best Local Similarity 57.1%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 HPGWGCGHPOG 17
I I I I I I I I
Db 39 HGGWGQPHGGG 52

RESULT 6

US-08-242-188-2
; Sequence 2, Application US/08242188
; Patent No. 5565186
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley B.
; APPLICANT: Scott, Michael R.
; TITLE OF INVENTION: METHOD OF DETECTING PRIONS IN A SAMPLE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Karl Bosicevic
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/242,188
FILING DATE: 13-MAY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Bosicevic, Karl
REGISTRATION NUMBER: 28,807
REFERENCE/DOCKET NUMBER: 06510/014001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 854-5277
TELEFAX: (415) 854-0875
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: HUMAN PRION PROTEIN, HuPrP

US-08-242-188-2
Query Match 43.1% Score 50; DB 1; Length 253;
Best Local Similarity 57.1%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 HPGWGCGHPOG 17
I I I I I I I I
Db 61 HGGWGQPHGGG 74

RESULT 7

US-08-509-261A-2
; Sequence 2, Application US/08509261A
; Patent No. 5763244
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley B.

; APPLICANT: Scott, Michael R.
; APPLICANT: Telling, Glenn
; TITLE OF INVENTION: Method of Detecting Prions
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bosicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/509,261A
FILING DATE: 31-JUL-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Bosicevic, Karl
REGISTRATION NUMBER: 28,807
REFERENCE/DOCKET NUMBER: 6510-030001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3400
TELEFAX: 650 327-3231
TELEX:

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-509-261A-2

Query Match 43.1% Score 50; DB 1; Length 253;
Best Local Similarity 57.1%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 HPGWGCGHPOG 17
I I I I I I I I
Db 61 HGGWGQPHGGG 74

RESULT 8

US-08-660-626-8
; Sequence 8, Application US/08660626
; Patent No. 5789655

GENERAL INFORMATION:
APPLICANT: Stanley B. Prusiner
APPLICANT: Glenn C. Telling
APPLICANT: Fred E. Cohen
APPLICANT: Michael R. Scott
TITLE OF INVENTION: TRANSGENIC ANIMALS EXPRESSING
TITLE OF INVENTION: EPITOPE-TAGGED PROTEINS
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson
STREET: 2200 Sand Hill Road, Suite 100
CITY: Menlo Park
STATE: California
COUNTRY: USA
ZIP: 94025

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

;
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/660,626
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Valeta Gregg
; REGISTRATION NUMBER: 35,127
; REFERENCE/DOCKET NUMBER: 07532/003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 322-5070
; TELEFAX: (415) 854-0875
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: HUMAN PRION PROTEIN, HuPrP
US-08-660-626-8

Query Match 43.1%; Score 50; DB 1; Length 253;
Best Local Similarity 57.1%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 HPQGWGGHPQGWG 17
| | | | |
Db 61 HGGGWGPHGGGWG 74

RESULT 9
US-08-692-892-2
; Sequence 2, Application US/08692892
; Patent No. 5792901
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley B.
; APPLICANT: Scott, Michael R.
; APPLICANT: Telling, Glenn
; TITLE OF INVENTION: DETECTING PRIONS IN A SAMPLE AND
; TITLE OF INVENTION: PRION PREPARATION AND TRANSGENIC ANIMAL USED FOR SAME
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Karl Bozicevic
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/692,892
; FILING DATE: 30-JULY-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bozicevic, Karl
; REGISTRATION NUMBER: 28,807
; REFERENCE/DOCKET NUMBER: 06510/056001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 322-5070
; TELEFAX: (415) 854-0875
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

;
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: HUMAN PRION PROTEIN, HuPrP
US-08-692-892-2

Query Match 43.1%; Score 50; DB 1; Length 253;
Best Local Similarity 57.1%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 HPQGWGGHPQGWG 17
| | | | |
Db 61 HGGGWGPHGGGWG 74

RESULT 10
US-08-713-939A-2
; Sequence 2, Application US/08713939A
; Patent No. 5846533
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley B.
; APPLICANT: Williamson, R. Anthony
; APPLICANT: Burton, Dennis R.
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR NATIVE PrP
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/713,939A
; FILING DATE: 13-SEP-1996
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bozicevic, Karl
; REGISTRATION NUMBER: 28,807
; REFERENCE/DOCKET NUMBER: 06510/059001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-5277
; TELEFAX: 415-854-0875
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-713-939A-2

Query Match 43.1%; Score 50; DB 2; Length 253;
Best Local Similarity 57.1%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 HPQGWGGHPQGWG 17
| | | | |
Db 61 HGGGWGPHGGGWG 74

RESULT 11
US-08-868-162A-22

; Sequence 22, Application US/08868162A
; Patent No. 5962669
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley
; APPLICANT: Cohen, Fred
; APPLICANT: James, Thomas
; APPLICANT: Kaneko, Kiyotoshi
; TITLE OF INVENTION: Prion Protein Modulator Factor
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/868,162A
; FILING DATE: 03-JUN-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bozicevic, Karl
; REGISTRATION NUMBER: 28,807
; REFERENCE/DOCKET NUMBER: 6510-083001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650-327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: HUMAN PRION PROTEIN, HuPrp
US-08-868-162A-22

Query Match 43.1%; Score 50; DB 2; Length 253;
Best Local Similarity 57.1%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 HPQGWGGHPQGW 17
| | | | |
Db 61 HGGWGPQHPGGGW 74

RESULT 12
US-09-031-168-8
; Sequence 8, Application US/09031168
; Patent No. 6150583
; GENERAL INFORMATION:
; APPLICANT: Stanley B. Prusiner
; APPLICANT: Glenn C. Telling
; APPLICANT: Fred E. Cohen
; APPLICANT: Michael R. Scott
; TITLE OF INVENTION: TRANSGENIC ANIMALS EXPRESSING
; TITLE OF INVENTION: EPITOPE-TAGGED PROTEINS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road, Suite 100

; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,168
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/660,626
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Valeta Gregg
; REGISTRATION NUMBER: 35,127
; REFERENCE/DOCKET NUMBER: 07532/003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 322-5070
; TELEFAX: (415) 854-0875
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: HUMAN PRION PROTEIN, HuPrp
US-09-031-168-8

Query Match 43.1%; Score 50; DB 4; Length 253;
Best Local Similarity 57.1%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 HPQGWGGHPQGW 17
| | | | |
Db 61 HGGWGPQHPGGGW 74

RESULT 13
US-09-128-450-20
; Sequence 20, Application US/09128450
; Patent No. 6211149
; GENERAL INFORMATION:
; APPLICANT: Chesebro, Bruce W
; APPLICANT: Caughey, Byron W
; APPLICANT: Chabry, Joelle
; APPLICANT: Priole, Susette
; TITLE OF INVENTION: Inhibitors of Formation of Protease Resistant Prion
; TITLE OF INVENTION: Protein
; FILE REFERENCE: 50121
; CURRENT APPLICATION NUMBER: US/09/128,450
; CURRENT FILING DATE: 1998-08-03
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-128-450-20

Query Match 43.1%; Score 50; DB 4; Length 253;
Best Local Similarity 57.1%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 HPQGWGGHPQGW 17
| | | | |

Db 61 HGGGQGPHGGGWG 74

RESULT 14

US-09-036-579-2
; Sequence 2, Application US/09036579
; Patent No. 6290954
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley B.
; APPLICANT: Williamson, R. Anthony
; APPLICANT: Burton, Dennis R.
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR NATIVE PrP
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036.579
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/713,939
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bozicevic, Karl
; REGISTRATION NUMBER: 28,807
; REFERENCE/DOCKET NUMBER: 06510/059001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-5277
; TELEFAX: 415-854-0875
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-036-579-2

Query Match 43.1%; Score 50; DB 4; Length 253;
Best Local Similarity 57.1%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 4 HPQGWGGHPOGWG 17
| | | | | | | |
Db 61 HGGGQGPHGGGWG 74

RESULT 15

US-09-823-494-20
; Sequence 20, Application US/09823494
; Patent No. 6355610
; GENERAL INFORMATION:
; APPLICANT: Chesebro, Bruce W
; APPLICANT: Caughey, Byron W
; APPLICANT: Chabry, Joelle
; APPLICANT: Priola, Susette
; TITLE OF INVENTION: Inhibitors of Formation of Protease Resistant Prion
; FILE REFERENCE: 50121
; CURRENT APPLICATION NUMBER: US/09/823,494
; CURRENT FILING DATE: 2001-03-30

; PRIOR APPLICATION NUMBER: 09/128,450
; PRIOR FILING DATE: 1998-08-03
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-823-494-20

Query Match 43.1%; Score 50; DB 4; Length 253;
Best Local Similarity 57.1%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 4 HPQGWGGHPOGWG 17
| | | | | | | |
Db 61 HGGGQGPHGGGWG 74

Search completed: August 6, 2002, 10:39:22
Job time: 159 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 10:37:30 ; Search time 35.27 Seconds
(without alignments)
46.315 Million cell updates/sec

Title: US-09-543-188A-34
Perfect score: 116
Sequence: 1 DGGHPQGGGGHPQGWG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
1	62	53.4	136 2 T29282	hypothetical prote
2	61.5	53.0	129 2 T21667	hypothetical prote
3	60	51.7	184 2 T16044	hypothetical prote
4	60	51.7	693 2 F96037	alpha-galactoside
5	60	51.7	695 2 A13129	hypothetical prote
6	60	51.7	727 2 H98157	periplasmic alpha-
7	59	50.9	96 2 S35715	glycine-rich prote
8	57	49.1	350 2 T36611	probable integral
9	56.5	48.7	1001 2 T17365	serine/threonine p
10	56	48.3	729 2 T51896	probable translati
11	54	46.6	104 2 JC4190	holotricin 3 precu
12	54	46.6	159 2 C49773	ecdysone-dependent
13	53	45.7	302 2 C84470	hypothetical prote
14	53	45.7	481 2 A26483	bindin precursor -
15	52.5	45.3	751 2 F87789	protein C34G6.2 [1
16	52	44.8	166 2 H96940	hypothetical prote
17	52	44.8	557 2 T47665	beta-N-acetylhexos
18	52	44.8	622 2 T37984	keratin 9, type I,
19	51.5	44.4	178 1 DDEC	single-stranded DN
20	51.5	44.4	178 2 A91259	ssDNA-binding prot
21	51.5	44.4	178 2 E86099	adhesive protein -
22	51	44.0	85 2 S42672	probable prephenat
23	51	44.0	327 2 B87218	Nifs-related amino
24	51	44.0	371 2 B86540	iron-sulfur cofact
25	51	44.0	371 2 B72085	cyclodiene insecti
26	51	44.0	533 2 S33744	hypothetical prote
27	51	44.0	846 2 H70599	fatty-acid synthas
28	51	44.0	2796 2 JC4743	protein R74.2 [imp
29	50.5	43.5	380 2 E88421	

ALIGNMENTS

RESULT 1

T29282

hypothetical protein C34D4.11 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000

C:Accession: T29282

R:Du, Z.; Le, T.T.

submitted to the EMBL Data Library, May 1996

A:Description: The sequence of C. elegans cosmid C34D4.

A:Reference number: Z20600

A:Accession: T29282

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-136 <DUZ>

A:Cross-references: EMBL:U58755; PIDN:AAB00596.1; GSPDB:GN00022; CESP:C34D4.11

A:Experimental source: strain Bristol N2; clone C34D4

C:Genetics:

A:Gene: CESP:C34D4.11

A:Map position: 4

A:Introns: 20/1; 66/1; 98/1; 116/1

C:Superfamily: Arabidopsis glycine-rich protein 3

Query Match 53.4%; Score 62; DB 2; Length 136;
Best Local Similarity 62.5%; Preq. No. 0.54;
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0

Qy 2 GGHPQGGGGHPQGWG 17

Db 58 GGGGGGGGGGGGGGGG 73

RESULT 2

T21667

hypothetical protein F32G8.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T21667

R:McMurray, A.

submitted to the EMBL Data Library, May 1996

A:Reference number: Z19456

A:Accession: T21667

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-129 <WIL>

A:Cross-references: EMBL:Z72509; PIDN:CAA96647.1; GSPDB:GN00023; CESP:F32G8.3

A:Experimental source: clone F32G8

C:Genetics:

A:Gene: CESP:F32G8.3

A:Map position: 5

A:Introns: 22/3; 73/3

Query Match 53.0%; Score 61.5; DB 2; Length 129;
Best Local Similarity 73.3%; Pred. No. 0.6;
Matches 11; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 2 GGHPOGWGGGHPOGW 16
||| ||||| |||
DB 33 GGGPGWGGG-PGW 46

RESULT 3
TI6044
hypothetical protein F12A10.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16044
R:Geisel, C.
submitted to the EMBL Data Library, June 1995
A:Description: The sequence of C. elegans cosmid F12A10.
A:Reference number: Z18451
A:Accession: T16044
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-184 <GE>
A:Cross-references: EMBL:U28731; NID:9861241; PID:9861242; PIDN:AAA68294.1; CESP:F12A10.
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:F12A10.1
A:Introns: 25/2; 56/1; 114/2

Query Match 51.7%; Score 60; DB 2; Length 184;
Best Local Similarity 60.0%; Pred. No. 1.3;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGHPOGWGGGHPOGW 16
||| ||||| |||
DB 125 GGGPGWGGG-PGW 139

RESULT 4
F96037
alpha-galactoside ABC transporter periplasmic solute-binding protein precursor agpA [imp
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: F96037
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1,683-Kb pSymB megaplasmid from the N2-fixing endo
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: F96037
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-693 <KUR>
A:Cross-references: GB:AL591985; PIDN:CAC49966.1; PID:gl5141454; GSPDB:GN00167
A:Experimental source: strain 1021, megaplasmid pSymB
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hynan, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: agpA; SMB21647
A:Genome: plasmid

Query Match 51.7%; Score 60; DB 2; Length 693;
Best Local Similarity 55.6%; Pred. No. 4.3;
Matches 10; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 2 GGHPOGW--GGHPOGW 17
||| ||| | |||
DB 111 GGRPEGWNYAGQTOGW 128

RESULT 5
AI3129
hypothetical protein agpA [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AI3129
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; MCCL
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AI3129
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-695 <KUR>
A:Cross-references: GB:AE008689; PIDN:AAL45455.1; PID:gl7743159; GSPDB:GN00187
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: agpA
A:Map position: linear chromosome

Query Match 51.7%; Score 60; DB 2; Length 695;
Best Local Similarity 55.6%; Pred. No. 4.3;
Matches 10; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 2 GGHPOGW--GGHPOGW 17
||| ||| | |||
DB 112 GGRPEGWNYAGSQGW 129

RESULT 6
H98157
periplasmic alpha-galactoside binding protein precursor [imported] - Agrobacterium tu
C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C:Accession: H98157
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldm
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacte
A:Reference number: A97359; PMID:11743194
A:Accession: H98157
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-727 <KUR>
A:Cross-references: GB:AE007870; PIDN:AAK88786.1; PID:gl5158535; GSPDB:GN00170
C:Genetics:
A:Gene: AGR_L_444
A:Map position: linear chromosome

Query Match 51.7%; Score 60; DB 2; Length 727;
Best Local Similarity 55.6%; Pred. No. 4.5;
Matches 10; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 2 GGHPOGW--GGHPOGW 17
||| ||| | |||
DB 144 GGRPEGWNYAGSQGW 161

RESULT 7
S35715
glycine-rich protein (clone DC 7.1), embryonic - carrot

```

Query Match          48.7%; Score 56.5; DB 2; Length 1001;
Best Local Similarity 57.9%; Pred. No.16;
Matches 11; Conservative 0; Mismatches 5; Indels 3; Gaps 1;

QY      2 GGHPQGWG---GGHPQGWG 17
        ||| ||| ||| ||| |||
Db       927 GGTPQAWGHPMQGGPQPWG 945

RESULT 10
T51896
Probable translation release factor erf3 [imported] - Neurospora crassa
N:Alternate names: protein B23I11.80
C:Species: Neurospora crassa
C>Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 03-Nov-2000
C:Accession: T51896
A:C.Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; NY
submitted to the Protein Sequence Database, August 2000
A:Reference number: Z25858
A:Accession: T51896
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-729 <SCH>
A:Cross-references: EMBL:AL391572; GSPDB:GN00116; NCSP:B23I11.80
A:Experimental source: BAC clone B23I11; strain OR74A
C:Genetics:
A:Gene: NCSP-B23I11.80
A:Map position: 6
A:Introns: 111/1; 711/3
C:Superfamily: suppressor 2 protein; translation elongation factor Tu homology

Query Match          48.3%; Score 56; DB 2; Length 729;
Best Local Similarity 55.0%; Pred. No.14;
Matches 11; Conservative 3; Mismatches 2; Indels 4; Gaps 1;

QY      2 GGHPQGWGGGHP----QGWG 17
        || | || | || | |||
Db       68 GGAQQCGGGYPQYGQQGYG 87

RESULT 11
JC4190
```

N:Alternate names: Antirungin protein
C:Species: Holotrichia diomphalia
C:Date: 04-Oct-1995 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999
C:Accession: JCA190
R:Lee, S.Y.; Moon, H.J.; Kurata, S.; Natori, S.; Lee, B.I.
Biol. Pharm. Bull. 18, 1049-1052, 1995
A:Title: Purification and cDNA cloning of an antifungal protein from the hemolymph of *Antirungia diomphalia*
A:Reference number: JCA190; MUID:96073722
A:Accession: JCA190
A:Molecule type: mRNA
A:Residues: 1-104 <LITE>
A:Cross-references: DDBJ: D13744; NID: g1088433; PIDN: BAA02889.1; PID: d1003394; P: P1003394
C:Comment: This protein is a Gly- and His-rich protein and a constitutive protein
C:Keywords: hemolymph
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-104/Product: holotriclin 3 #status predicted <MAY>

Query Match	46.6%;	Score 54;	DB 2;	Length 104;
Best Local Similarity	62.5%;	Pred. No. 4;		
Matches 10;	Conservative 0;	Mismatches 6;	Indels 0;	Gaps 0;

QY 2 GGHPCGGGGCHPCGW 17
| | | | | | | |
Db 29 GGHGGGGGGHGMGQG 44

RESULT 12
C49773

ecdysone-dependent cytochrome-like protein EDG-91 - fruit fly (*Drosophila melanogaster*)
C:Species: *Drosophila melanogaster*
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 24-Nov-1999
C:Accession: C49773
R:Apple, R.T.; Fristrom, J.W.
Dev. Biol. 146, 569-582, 1991
A:Title: 20-Hydroxyecdysone is required for, and negatively regulates, transcription of
A:Reference number: A49773; MUID:91323677
A:Accession: C49773
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-159 <APP>
A:Cross-references: GB:M1250; NID:g157326; PID:g157327; GB:M1251; NID:g157328; PID:g157329
C:Genetics:
A:Gene: FlyBase:Edg91
A:Cross-references: FlyBase:FBgn0004554
C:Superfamily: glycine-rich cell wall structural protein 1

Query Match 46.6%; Score 54; DB 2; Length 159;
Best Local Similarity 50.0%; Pred. No. 5.9;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 2 GGHPOGCGGHPQGWG 17
||| | | | | |
Db 61 GGYPGYASGYPGGY 76

RESULT 13
C84470
hypothetical protein At2g05580 [imported] - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: C84470
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.A.; VanKen, S.E.; Unayam, L.; Tallon, L.;
Euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
A:Reference number: A84420; MUID:20083487
A:Accession: C84470
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-302 <STO>
A:Cross-references: GB:AE002093; NID:g4581166; PIDN:AAD24649.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g05580
A:Map position: 2

Query Match 45.7%; Score 53; DB 2; Length 302;
Best Local Similarity 62.5%; Pred. No. 14;
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 2 GGHPOGCGGHPQGWG 17
||| | | | | |
Db 256 GGHKGCGGGGHHVGGG 271

RESULT 14
A26483
bindin precursor - sea urchin (*Strongylocentrotus purpuratus*)
C:Species: *Strongylocentrotus purpuratus* (purple urchin)
C:Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 20-Mar-1998
C:Accession: A26483
R:Gao, B.; Klein, L.E.; Britten, R.J.; Davidson, E.H.
Proc. Natl. Acad. Sci. U.S.A. 83, 8634-8638, 1986
A:Title: Sequence of mRNA coding for bindin, a species-specific sea urchin sperm protein
A:Reference number: A26483; MUID:87041508
A:Accession: A26483
A:Molecule type: mRNA
A:Residues: 1-481 <GAO>
A:Cross-references: GB:M14487; NID:g161444; PID:g161445

C:Keywords: sperm

Query Match 45.7%; Score 53; DB 2; Length 481;
Best Local Similarity 54.5%; Pred. No. 22;
Matches 12; Conservative 0; Mismatches 4; Indels 6; Gaps 1;
QY 2 GGHPOGCGGHPQGWG 17
||| | | | | |
Db 445 GGPPOGCGGHPQGWG 466

RESULT 15
F87789
protein C34G6.2 [imported] - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: F87789
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating bio
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Project/ -
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
A:Accession: F87789
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-751 <STO>
A:Cross-references: GB:chr_I; PIDN:AAB52481.1; PID:g1943843; GSPDB:GN00019; CESP:C34G
C:Genetics:
A:Gene: C34G6.2
A:Map position: 1

Query Match 45.3%; Score 52.5; DB 2; Length 751;
Best Local Similarity 62.5%; Pred. No. 38;
Matches 10; Conservative 0; Mismatches 3; Indels 3; Gaps 1;
QY 2 GGHPOGCGGHPQGWG 17
||| | | | | |
Db 713 GGGRGCGGG---GWG 725

Search completed: August 6, 2002, 10:37:32
Job time: 49 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 10:40:51 ; Search time 18.57 Seconds
(without alignments)
35.446 Million cell updates/sec

Title: US-09-543-188a-34

Perfect score: 116

Sequence: 1 DGGHPQGMGGHPQGMG 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match %	Length	ID	Description
1	60	51.7	693	1 AGPA_RHIME	O9x4yl rhizobium m
2	59	50.9	96	1 GRP7_DAUCA	P37704 daucus caro
3	58	50.0	259	1 PRIO_TRIVU	P51780 trichosurus
4	54	46.6	104	1 HOL3_HOLDI	Q25055 holotrichia
5	54	46.6	159	1 CUP9_DROME	P27781 drosophila
6	53	45.7	481	1 BIND_STRPU	P06651 strongyloce
7	52	44.8	256	1 PRP2_TRAST	P40243 tragelephus
8	52	44.8	622	1 K1CI_HUMAN	P35527 homo sapien
9	51.5	44.4	177	1 SSB_ECOLI	P02339 escherichia
10	50.5	43.5	102	1 SGP3_CHRVI	O52055 chromatium
11	50	43.1	232	1 PRIO_AVEGE	P40246 ateles geof
12	50	43.1	238	1 PRIO_CERAT	Q95145 cercocobus
13	50	43.1	238	1 PRIO_THERGE	Q95270 thetopithec
14	50	43.1	239	1 PRIO_AOTTR	P40245 aotus trivi
15	50	43.1	241	1 PRIO_CALMO	P40248 callicebus
16	50	43.1	241	1 PRIO_MANSP	P40255 mandrillus
17	50	43.1	245	1 PRIO_CERAE	P40250 cercopithec
18	50	43.1	246	1 PRIO_CERMO	Q95172 cercopithec
19	50	43.1	246	1 PRIO_CERPA	Q95174 cercopithec
20	50	43.1	246	1 PRIO_CERTO	Q95176 cercocobus
21	50	43.1	252	1 PRIO_ATEPA	P51446 ateles pani
22	50	43.1	252	1 PRIO_CALJA	P40247 callithrix
23	50	43.1	252	1 PRIO_CEBAP	P40249 cebus apell
24	50	43.1	252	1 PRIO_RABIT	Q95211 oryctolagus
25	50	43.1	253	1 PRIO_COLGU	P40251 colobus gue
26	50	43.1	253	1 PRIO_GORGO	P40252 gorilla gor
27	50	43.1	253	1 PRIO_HUMAN	P04156 homo sapien
28	50	43.1	253	1 PRIO_MACFA	P40254 macaca fasc
29	50	43.1	253	1 PRIO_PANTR	P40253 pan troglod
30	50	43.1	253	1 PRIO_PONPY	P40256 pongo pygma
31	50	43.1	253	1 PRIO_PREFR	P40257 presbytis f
32	50	43.1	254	1 PRIO_CRIGR	Q60506 cricetus
33	50	43.1	254	1 PRIO_CRIMI	Q60468 cricetus

P04273 mesocricetu
P13852 rattus norv
Q92053 sigmodon hi
P79141 camelus dro
O46501 canis fami
P52113 capra hircu
P79142 cervus elap
O18754 felis silve
P47852 odocoileus
P23907 ovis aries
Q01880 bos taurus
P52114 mustela put

ALIGNMENTS

RESULT 1
AGPA_RHIME 1
ID AGPA_RHIME STANDARD; PRT; 693 AA.
AC Q9X4Y1;
DT 30-MAY-2000 (Rel. 39, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Periplasmic alpha-galactoside binding protein precursor.
GN AGPA OR RB1567 OR SMB21647.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymb (megaplasmid 2).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
ON NCBI_TaxID=382;
RX STRAIN=1021.
RX MEDLINE=99009011; PubMed=9791127;
RA Gage D.J., Long S.R.;
RT "Alpha-galactoside uptake in Rhizobium meliloti: isolation and characterization of agpA, a gene encoding a periplasmic binding protein required for melibiose and raffinose utilization.";
RL J. Bacteriol. 180:5739-5748(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396508; PubMed=11481431;
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P., Vorheeler F.J., Hernandez-Lucas I., Becker A., Gouzy J., Golding B., Puehler A.;
RT "The complete sequence of the 1.683-kb pSymb megaplasmid from the N2-fixing endosymbiont Sinorhizobium meliloti.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
CC -1- FUNCTION: PROBABLY INVOLVED IN AN ALPHA-GALACTOSIDES INTAKE TRANSPORT SYSTEM.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING PROTEIN FAMILY 5.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF119834; AAD26274.1; -
CC EMBL; AL603647; CAC49966.1; -
CC InterPro; IPR000914; SBP_bac_5.
CC Pfam; PF00496; SBP_bac_5; 2.
CC PROSITE; PS01040; SBP_BACTERIAL_5; FALSE_NEG.
CC Sugar transport; transpor; Periplasmic; Signal; Plasmid;
CC Complete proteome.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 693 PERIPLASMIC ALPHA-GALACTOSIDE BINDING

```
FT CONFLICT 397 397 PROTEIN.  
SQ SEQUENCE 693 AA: 77407 MW; 76F9B95708C2DF9F CRC64;  
  
Query Match 51.7%; Score 60; DB 1; Length 693;  
Best Local Similarity 55.6%; Pred. No. 1.4;  
Matches 10; Conservative 1; Mismatches 5; Indels 2; Gaps 1;  
  
Oy 2 GCGHPOGW--GGCHPOGW 17  
Db 111 GGRPEGWNSAGQTQGW 128  
  
RESULT 2  
GRP7_DAUCA  
ID GRP7_DAUCA STANDARD; PRT; 96 AA.  
AC P37704;  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DE Glycine-rich protein DC7.1 precursor.  
OS Daucus carota (Carrot).  
CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
CC Asteridae; euasterids II; Apiales; Apiaceae; Daucus.  
CC NCBI_TaxID=4039;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Aleith F., Richter G.;  
RT "Gene expression during induction of somatic embryogenesis in carrot  
cell suspensions.";  
RL Planta 183:17-24(1990).  
CC -1- FUNCTION: MAY BE CONNECTED WITH THE INITIATION OF EMBRYOGENESIS  
CC OR WITH THE METABOLIC CHANGES PRODUCED BY THE REMOVAL OF AUXINS.  
CC -1- DEVELOPMENTAL STAGE: TRANSIENTLY EXPRESSED DURING EARLY  
CC EMBRYOGENESIS.  
CC -1- INDUCTION: BY THE REMOVAL OF AUXINS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC EMBL; X15706; CAA33736.1; -.  
DR PIR; S35715; S35715.  
KW Repeat; Signal.  
FT SIGNAL 1 25 POTENTIAL.  
FT CHAIN 26 96 GLYCINE-RICH PROTEIN DC7.1.  
FT DOMAIN 42 67 2 APPROXIMATE REPEATS OF H-H-G(4,6)-H.  
FT REPEAT 42 50 1.  
FT REPEAT 61 67 2.  
SQ SEQUENCE 96 AA: 9319 MW; 7C00D44637B7A364 CRC64;  
  
Query Match 50.9%; Score 59; DB 1; Length 96;  
Best Local Similarity 64.7%; Pred. No. 0.31;  
Matches 11; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
  
Oy 1 DGGHPQGGGGHPOGW 17  
Db 39 DGGHHGGGGGGHYSGG 55  
  
RESULT 3  
PRIO_TRIVU  
ID PRIO_TRIVU STANDARD; PRT; 259 AA.  
AC P51780;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)
```

```
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C).  
GN PRNP.  
OS Trichosurus vulpecula (Brush-tailed possum).  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Metazoa; Diprotodontia; Phalangeridae; Trichosurus.  
CC NCBI_TaxID=9337;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Windl O., Dempster M., Estibeiro P., Lathe R.;  
RL Submitted (May-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE  
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.  
CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED  
CC "RODS".  
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.  
CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND  
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,  
CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME  
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),  
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.  
CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC EMBL; L38993; AAA61833.1; -.  
DR HSP; P04925; IAG2.  
DR InterPro: IPR000817; Prion.  
DR Pfam: PF00377; prion; 1.  
DR PRINTS; PR00341; PRION.  
DR SMART; SM00157; PRP; 1.  
DR PROSITE; PS00291; PRION_1; 1.  
DR PROSITE; PS00706; PRION_2; 1.  
KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.  
FT SIGNAL 1 24 BY SIMILARITY.  
FT CHAIN 25 259 MAJOR PRION PROTEIN.  
FT DISULFID 184 219 BY SIMILARITY.  
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 202 202 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 259 AA: 28550 MW; 27DA7D8EBCA5C892 CRC64;  
  
Query Match 50.0%; Score 58; DB 1; Length 259;  
Best Local Similarity 57.1%; Pred. No. 1;  
Matches 12; Conservative 0; Mismatches 3; Indels 6; Gaps 2;  
  
Oy 3 GHPOG---WGCGHPOG--WG 17  
Db 59 GHPOGGGTNWGPHPGGSNWG 79  
  
RESULT 4  
HOL3_HOLDI  
ID HOL3_HOLDI STANDARD; PRT; 104 AA.  
AC Q25055;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Holotricin 3 precursor.  
OS Holotrichia diomphalia.  
CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
CC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;  
CC Scarabaeiformia; Scarabaeidae; Melolonthinae; Holotrichia.  
CC NCBI_TaxID=33394;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 21-40.  
RC TISSUE=Larval hemolymph;
```

```

RX MEDLINE=96073722; PubMed=8535393;
RA Lee S.Y., Moon H.J., Kurata S., Natori S., Lee B.L.;
RT "Purification and cDNA cloning of an antifungal protein from the
RL hemolymph of Holotrichia diomphalia larvae.";
CC Biol. Pharm. Bull. 18:1049-1052(1995).
CC
CC -!- FUNCTION: HOLOTRICIN 3 HAS ANTIFUNGAL ACTIVITY AGAINST
CC C. ALBICANS.
CC
CC -!- SIMILARITY: TO TENECIN 3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D13744; BRA02889.1; -
DR HSSP; P30129; 4DPV.
KW Insect immunity; Antibiotic; Hemolymph; Fungicide; Signal; Repeat.
FT SIGNAL 1 20
FT PEPTIDE 21 104
FT DOMAIN 27 98
FT REPEAT 27 30
FT REPEAT 31 34
FT REPEAT 35 38
FT REPEAT 39 42
FT REPEAT 43 46
FT REPEAT 47 50
FT REPEAT 51 54
FT REPEAT 55 58
FT REPEAT 59 62
FT REPEAT 63 66
FT REPEAT 67 70
FT REPEAT 71 74
FT REPEAT 75 78
FT REPEAT 79 82
FT REPEAT 83 86
FT REPEAT 87 90
FT REPEAT 91 94
FT REPEAT 96 98
SQ SEQUENCE 104 AA; 9036 MW; 2799D681BFDCC725 CRC64;

Query Match 46.6%; Score 54; DB 1; Length 104;
Best Local Similarity 62.5%; Pred. No. 1.4;
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 GGHPOGGGGHPOGGW 17
Db 29 GGHGGGGHGGHGGG 44

RESULT 5
CUP9_DROME STANDARD; PRT; 159 AA.
AC P27781; OSVE11;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Pupal cuticle protein Edg-91 precursor (Ecdysone-dependent protein
DE 91).
GN EDG91 OR EDG-91 OR CG7539.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=91323677; PubMed=1713868;

```

```

RA Apple R.T., Fristrom J.W.;
RT "20-Hydroxyecdysone is required for, and negatively regulates,
RT transcription of Drosophila pupal cuticle protein genes.";
RL Dev. Biol. 146:569-582(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balley R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lal Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J.R., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC
CC -!- FUNCTION: COMPONENT OF THE PUPAL CUTICLE.
CC -!- TISSUE SPECIFICITY: LARVAL (POSTERIOR) AND IMAGINAL (ANTERIOR)
CC EPIDERMIS.
CC
CC -!- DEVELOPMENTAL STAGE: PREPUPAL AND EARLY PUPAL STAGES.
CC -!- DOMAIN: THIS PROTEIN IS GLYCINE-RICH AND CONTAINS SEVERAL REPEATS
CC OF THE MOTIF (C/S)1-4(Y/F) LIKE STRUCTURAL PROTEINS FROM INSECT
CC EGG SHELLS, EGG CASES AND VERTEBRATE CYTOKERATINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M71250; AAA28502.1; -
DR EMBL; M71251; AAA28503.1; -
DR EMBL; AE003718; AAF55440.1; -
DR PIR; C49773; C49773.
DR FlyBase; FBgn0004554; Edg91.
KW Structural protein; Cuticle; Signal.
FT SIGNAL 1 21
FT CHAIN 22 159 PUPAL CUTICLE PROTEIN EDG-91.
FT DOMAIN 35 159 GLY/TYR-RICH.
FT CONFLICT 29 109 S -> R (IN REF. 1).
FT CONFLICT 109 109 H -> Y (IN REF. 1).

```

445 GGPPQGMGMPPHQGMGGPPQGMG 46

RESULT 8
 K1C1_HUMAN
 ID K1C1_HUMAN STANDARD; PRT; 622 AA.
 AC P35527; O14665;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Keratin, type I cytoskeletal 9 (Cytokeratin 9) (K9) (CK 9).
 GN KRT9.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RC TISSUE=Foot sole tissue; PubMed=7507869;
 RA Langbein L., Heid H.W., Moll I., Franke W.W.;
 RX MEDLINE=9413120; PubMed=7507869;
 RA Langbein L., Heid H.W., Moll I., Franke W.W.;
 RT "Molecular characterization of the body site-specific human epidermal
 cytokeatin 9: cDNA cloning, amino acid sequence, and tissue
 specificity of gene expression.";
 RL Differentiation 55:57-72(1993).
 RN [2]
 RP SEQUENCE OF 449-465.
 RX MEDLINE=90267446; PubMed=2140676;
 RA Rosen E.M., Meromsky L., Romero R., Setter E., Goldberg I.;
 RT "Human placenta contains an epithelial scatter protein.";
 RL Biochem. Biophys. Res. Commun. 168:1082-1088(1990).
 RN [3]
 RP VARIANTS EPPK VAL-156 AND PRO-171.
 RX MEDLINE=94274199; PubMed=7516304;
 RA Hennies H.-C., Zehender D., Kunze J., Kuester W., Reis A.;
 RT "Keratin 9 gene mutational heterogeneity in patients with
 epidermolytic palmoplantar keratoderma.";
 RL Hum. Genet. 93:649-654(1994).
 RN [4]
 RP VARIANT EPPK SER-160.
 RX MEDLINE=95015968; PubMed=7523529;
 RA Bonifas J.M., Matsumura K., Chen M.A., Berth-Jones J.,
 RA Hutchinson P.E., Zloczower M., Fritsch P.O., Epstein E.H. Jr.;
 RT "Mutations of keratin 9 in two families with palmoplantar
 epidermolytic hyperkeratosis.";
 RL J. Invest. Dermatol. 103:474-477(1994).
 RN [5]
 RP VARIANT EPPK TYR-160.
 RX MEDLINE=94184355; PubMed=7511021;
 RA Torchard D., Blanchet-Bardon C., Serova O., Langbein L., Narod S.,
 RA Janin N., Goguel A.F., Bernheim A., Franke W.W., Lenoir G.M.,
 RA Feunteun J.;
 RT "Epidermolytic palmoplantar keratoderma cosegregates with a keratin 9
 mutation in a pedigree with breast and ovarian cancer.";
 RL Nat. Genet. 6:106-110(1994).
 RN [6]
 RP VARIANTS EPPK LYS-160; GLN-162 AND TRP-162.
 RX MEDLINE=94214498; PubMed=7512862;
 RA Reis A., Hennies H.-C., Langbein L., Digweed M., Mischke D.,
 RA Dreschler M., Schroek E., Royer-Pokora B., Franke W.W., Sperling K.,
 RA Kuester W.;
 RT "Keratin 9 gene mutations in epidermolytic palmoplantar keratoderma
 (EPPK).";
 RL Nat. Genet. 6:174-179(1994).
 RN [7]
 RP VARIANTS EPPK TRP-162 AND SER-167.
 RX MEDLINE=95164983; PubMed=7532199;
 RA Rothnagel J.A., Wojcik S., Liefer K.M., Dominey A.M., Huber M.,
 RA Hohl D., Roop D.R.;
 RT "Mutations in the 1A domain of keratin 9 in patients with
 epidermolytic palmoplantar keratoderma.";
 RL J. Invest. Dermatol. 104:430-433(1995).
 RN [8]
 RP VARIANT EPPK VAL-159.
 RX MEDLINE=97348961; PubMed=9204965;
 RA Endo H., Hatamochi A., Shinkai H.;

"A novel mutation of a leucine residue in coil 1A of keratin 9 in
 epidermolytic palmoplantar keratoderma.";
 J. Invest. Dermatol. 109:113-115(1997).
 RN [9]
 RP VARIANTS EPPK THR-156; VAL-156 AND GLN-162.
 RX MEDLINE=99072662; PubMed=9856842;
 RA Covello S.P., Irvine A.D., McKenna K.E., Munro C.S., Nevlin N.C.,
 RA Smith F.J.D., Uitto J., McLean W.H.I.;
 RT "Mutations in keratin 9 in kindreds with epidermolytic palmoplantar
 keratoderma and epidemiology in Northern Ireland.";
 RL J. Invest. Dermatol. 111:1207-1209(1998).
 CC -!- FUNCTION: MAY SERVE AN IMPORTANT SPECIAL FUNCTION EITHER IN THE
 MATURE PALMAR AND PLANTAR SKIN TISSUE OR IN THE MORPHOGENIC
 PROGRAM OF THE FORMATION OF THESE TISSUE.
 CC -!- SUBUNIT: HETROTETRAMER OF TWO TYPE I AND TWO TYPE II KERATINS.
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE TERMINALLY DIFFERENTIATED
 EPIDERMIS OF PALMS AND SOLES.
 CC -!- DISEASE: DEFECTS IN KRT9 ARE A CAUSE OF EPIDERMOLYTIC PALMOPLANTAR
 KERATODERMA (EPPK), AN AUTOSOMAL DOMINANT DISEASE CHARACTERIZED BY
 DIFFUSE THICKENING OF THE EPIDERMIS ON THE ENTIRE SURFACE OF
 PALMS AND SOLES SHARPLY BORDERED WITH ERYTHEMATOUS MARGINS.
 CC -!- MISCELLANEOUS: THERE ARE TWO TYPES OF CYTOSKELETAL AND
 MICROFIBRILLAR KERATIN: I (ACIDIC; 40-55 kDa) [K9 TO K20] AND II
 (NEUTRAL TO BASIC; 56-70 kDa) [K1 TO K8].
 CC -!- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -!- CAUTION: WAS ORIGINALLY (REF.2) THOUGHT TO BE A 60 kDa CHAIN OF
 PLACENTAL SCATTER PROTEIN

 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 CC EMBL; X75015; CAAS2924.1; -;
 DR EMBL; Z29074; CAAB2315.1; -;
 DR EMBL; S69510; AAC60619.1; -;
 DR PIR; B35494; B35494.
 DR HSP; P02876; 2WGC.
 DR MIM; 144200; -;
 DR InterPro; IPR001664; IF.
 DR InterPro; IPR002957; Keratin_I.
 DR Pfam; PF00038; filament; 1.
 DR PRINTS; PR01248; TYPE1KERATIN.
 DR PROSITE; PS00226; IF; 1.
 KW Intermediate filament; Coiled coil; Keratin; Disease mutation.
 FT DOMAIN 1 151
 FT HEAD.
 FT ROD.
 FT TAIL.
 FT COIL 1A.
 FT LINKER 1.
 FT COIL 1B.
 FT LINKER 12.
 FT COIL 2.
 FT POLY-GLY.
 FT M -> T (IN EPPK).
 FT /FTid=VAR_010499.
 FT M -> V (IN EPPK).
 FT /FTid=VAR_010500.
 FT L -> V (IN EPPK).
 FT /FTid=VAR_010501.
 FT N -> K (IN EPPK).
 FT /FTid=VAR_003822.
 FT N -> S (IN EPPK).
 FT /FTid=VAR_010502.
 FT N -> Y (IN EPPK).
 FT /FTid=VAR_010503.
 FT R -> Q (IN EPPK).
 FT /FTid=VAR_003823.
 FT R -> W (IN EPPK).
 FT /FTid=VAR_003824.
 FT VARIANT 156 156
 FT VARIANT 159 159
 FT VARIANT 160 160
 FT VARIANT 160 160
 FT VARIANT 160 160
 FT VARIANT 160 160
 FT VARIANT 162 162
 FT VARIANT 162 162

FT VARIAT 167 167 L -> S (IN EPPK).
FT FTID-VAR_003825.
FT VARIAT 171 171 Q -> P (IN EPPK).
FT FTID-VAR_010504.
FT CONFLICT 12 12 T -> SR (IN REF. 1: CAAB2315).
SQ SEQUENCE 622 AA; 61987 MW; 898C3825D4B5ED94 CRC64;

Query Match 44.8%; Score 52; DB 1; Length 622;
Best Local Similarity 62.5%; Pred. No. 13;
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GCHPOGCGGHPQGWG 17
||| | |||| | |
DB 539 GCHSGSGGGHSGSGG 554

RESULT 9
SSB_ECOLI STANDARD; PRT; 177 AA.
AC P02339;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Single-strand binding protein (SSB) (Helix-destabilizing protein).
GN SSB OR EXRB OR LEXC OR B4059 OR Z5658 OR ECS5041.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-52.
RX MEDLINE=82037821; PubMed=6270666;
RA Sancar A., Williams K.R., Chase J.W., Rupp W.D.;
RT "Sequences of the ssb gene and protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:4274-4278(1981).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=94089392; PubMed=8265357;
RA Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
RT Daniels D.L.;
RL "Analysis of the Escherichia coli genome. IV. DNA sequence of the
region from 89.2 to 92.8 minutes.";
RN Nucleic Acids Res. 21:5408-5417(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RL "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RN Nature 409:529-533(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kubara S., Shiba T., Hattori M., Shinagawa H.;
RL "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RN Nature Res. 8:11-22(2001).
RN [5]
RP SEQUENCE OF 1-40.
RX MEDLINE=82186711; PubMed=7042342;
RA Beyreuther K., Berthold-Schmidt V., Geider K.;

RT "Biological activity and a partial amino-acid sequence of Escherichia
coli DNA-binding protein I isolated from overproducing cells.";
RL Eur. J. Biochem. 123:415-420(1982).
RN [6]
RX CHARACTERIZATION, AND REVISION TO 133.
RX MEDLINE=84111559; PubMed=6363409;
RA Chase J.W., L'Italien J.J., Murphy J.B., Spicer E.K., Williams K.R.;
RT "Characterization of the Escherichia coli SSB-113 mutant
single-stranded DNA-binding protein. Cloning of the gene, DNA and
protein sequence analysis, high pressure liquid chromatography
peptide mapping, and DNA-binding studies.";
RL J. Biol. Chem. 259:805-814(1984).
RN [7]
RX MUTANT SSB-1.
RX MEDLINE=85006904; PubMed=6384214;
RA Williams K.R., Murphy J.B., Chase J.W.;
RT "Characterization of the structural and functional defect in the
Escherichia coli single-stranded DNA binding protein encoded by the
ssb-1 mutant gene. Expression of the ssb-1 gene under lambda pL
regulation.";
RL J. Biol. Chem. 259:11804-11811(1984).
RN [8]
RX MUTANT SSB-1, AND DNA-BINDING.
RX MEDLINE=91108818; PubMed=1988680;
RA Bujalowski W., Lohman T.M.;
RT "Monomers of the Escherichia coli SSB-1 mutant protein bind single-
stranded DNA.";
RL J. Mol. Biol. 217:63-74(1991).
RN [9]
RX MUTANTS SSB-200; SSB-201 AND SSB-202.
RC STRAIN=K12 / W3110;
RX MEDLINE=97284491; PubMed=9139905;
RA Reddy M., Gowrishankar J.;
RT "Identification and characterization of ssb and uup mutants with
increased frequency of precise excision of transposon Tn10
derivatives: nucleotide sequence of uup in Escherichia coli.";
RL J. Bacteriol. 179:2892-2899(1997).
RN [10]
RX MUTAGENESIS, AND DNA-BINDING.
RX MEDLINE=87276593; PubMed=3301414;
RA Casas-Finet J.R., Khamis M.I., Maki A.W., Chase J.W.;
RT "Tryptophan 54 and phenylalanine 60 are involved synergistically in
the binding of E. coli SSB protein to single-stranded
polynucleotides.";
RL FEBS Lett. 220:347-352(1987).
RN [11]
RX REVIEW
RX MEDLINE=91203755; PubMed=2087220;
RA Meyer R.R., Laine P.S.;
RT "The single-stranded DNA-binding protein of Escherichia coli.";
RL Microbiol. Rev. 54:342-380(1990).
RN [12]
RX X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 1-135.
RX MEDLINE=97338073; PubMed=9192620;
RA Raghunathan S., Ricard C.S., Lohman T.M., Waksman G.;
RT "Crystal structure of the homo-tetrameric DNA binding domain of
Escherichia coli single-stranded DNA-binding protein determined by
multiwavelength X-ray diffraction on the selenomethionyl protein at
2.9-A resolution.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:6652-6657(1997).
RN [13]
RX X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 1-135.
RX MEDLINE=20392541; PubMed=10932248;
RA Raghunathan S., Kozlov A.G., Lohman T.M., Waksman G.;
RT "Structure of the DNA binding domain of E. coli SSB bound to ssDNA.";
RL Nat. Struct. Biol. 7:648-652(2000).
CC -1- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR REPLICATION OF THE
CHROMOSOMES AND ITS SINGLE-STRANDED DNA PHAGES. IT IS ALSO
INVOLVED IN DNA RECOMBINATION AND REPAIR.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SIMILARITY: BELONGS TO THE SSB FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; J01704; AAA24649.1; -
 EMBL; U00006; AAC43153.1; -
 EMBL; AE000479; AAC77029.1; -
 EMBL; AE005639; AAG59257.1; -
 EMBL; AP002568; BAB38464.1; -
 PIR; A02681; DDEC.
 PDB; 1KAW; 31-DEC-97.
 PDB; 1EYG; 01-AUG-00.
 SWISS-2DPAGE; P02339; COLI.
 ECO2DBASE; F018.8; 6TH EDITION.
 Ecogene; EG10976; ssb.
 InterPro; IPR000424; SSB.
 Pfam; PF00436; SSB; 1.
 PROSITE; PS00735; SSB.1; 1.
 PROSITE; PS00736; SSB.2; 1.
 DNA-binding; DNA repair; DNA replication; 3D-structure;
 Complete proteome.
 INIT_MET 0 0
 DNA_BIND 54 60
 MUTAGEN 4 4
 G->D: INCREASED FREQUENCY OF PRECISE
 EXCISION OF TRANSPOSON TN10 DERIVATIVES
 (MUTANT SSB-200).
 L->F: INCREASED FREQUENCY OF PRECISE
 EXCISION OF TRANSPOSON TN10 DERIVATIVES
 (MUTANT SSB-202).
 P->S: INCREASED FREQUENCY OF PRECISE
 EXCISION OF TRANSPOSON TN10 DERIVATIVES
 (MUTANT SSB-202).
 H->I: DESTABILIZES THE TETRAMER (MUTANT
 SSB-1).
 H->L: REDUCES DNA-BINDING AFFINITY.
 F->A: REDUCES DNA-BINDING AFFINITY.
 V->M: INCREASED FREQUENCY OF PRECISE
 EXCISION OF TRANSPOSON TN10 DERIVATIVES
 (MUTANT SSB-201).
 G -> S (IN REF. 1).
 CONFLICT 133 133
 SEQUENCE 177 AA; 18844 MW; 8D9B1E346E46A0B5 CRC64;

Query Match 44.4%; Score 51.5; DB 1; Length 177;
 Best Local Similarity 63.2%; Pred. No. 4.6;
 Matches 12; Conservative 0; Mismatches 4; Indels 3; Gaps 2;

QY 2 GGHPQGW--GGHPQ-GWG 17
 ||| | ||| ||| |||
 Db 118 GGAPAGNIGGGQPGQGW 136

RESULT 10
 SGP3_CHRV1 STANDARD; PRT; 102 AA.
 ID OS2055;
 AC 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Sulfur globule protein CV3 precursor.
 GN SGP3 OR SGP3
 OS Chromatium vinosum.
 OC Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae;
 OC Allochrochium.
 OX NCBI_TaxID=1049;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-D / ATCC 17899 / DSM 180;
 RX MEDLINE=98228197; PubMed=9560425;
 RA Pattaragulwanit K., Brune D.C., Trueper H.G., Dahl C.;

Molecular genetic evidence for extracytoplasmic localization of sulfur globules in Chromatium vinosum.;
 Arch. Microbiol. 169:434-444(1998).
 [2]
 RC SEQUENCE OF 26-78, AND CHARACTERIZATION.
 RP STRAIN-D / ATCC 17899 / DSM 180;
 RX MEDLINE=96009136; PubMed=7575095;
 RA Brune D.C.;
 RT *Isolation and characterization of sulfur globule proteins from Chromatium vinosum and Thiocapsa roseopersicina.*;
 Arch. Microbiol. 163:391-399(1995).
 CC -1- FUNCTION: STRUCTURAL PROTEIN OF THE SULFUR GLOBULES, WHICH ARE INTRACELLULAR GLOBULES THAT SERVE FOR SULFUR STORAGE IN PURPLE SULFUR BACTERIA.
 CC -1- SUBUNIT: THE PROTEIN ENVELOPE OF THE SULFUR GLOBULES ARE COMPOSED OF THREE DIFFERENT PROTEINS; CV1, CV2 AND CV3.
 CC -1- MASS SPECTROMETRY: MW=8479; METHOD=WALDI; RANGE=26-102.
 CC -1- SIMILARITY: TO T.ROSEPERSICINA TR2.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; AF017118; AAB91542.1; -
 Structural protein; Signal.
 FT SIGNAL 1 25
 FT CHAIN 26 102 Sulfur globule protein CV3.
 SQ SEQUENCE 102 AA; 10979 MW; 02E461F2270D6B23 CRC64;

Query Match 43.5%; Score 50.5; DB 1; Length 102;
 Best Local Similarity 58.8%; Pred. No. 3.7;
 Matches 10; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 2 GGHPQGWGGGHPQ-GWG 17
 | : | | | | | | | | | |
 Db 28 GWNFGWGGGHPDGPWG 44

RESULT 11
 PRT_ATEGE STANDARD; PRT; 232 AA.
 ID PA0246;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C) (Fragment).
 GN PRNP
 OS Ateles geoffroyi (Black-handed spider monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateles.
 OX NCBI_TaxID=9509;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95139066; PubMed=7837269;
 RA Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;
 RT "Prion protein gene variation among primates.";
 RL J. Mol. Biol. 245:362-374(1995).
 CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
 CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED "RODS".
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU, CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE), TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
 CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; U08309; AAC50097.1; -
 DR HSSP; P04156; 1E1G
 DR InterPro; IPR000817; Prion.
 DR Pfam; PF00377; prion; 1.
 DR SMART; SM00157; PRP; 1.
 DR PROSITE; PS00291; PRION_1; 1.
 DR PROSITE; PS00706; PRION_2; 1.
 KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
 FT NON_TER 1
 FT SIGNAL <1 15
 FT CHAIN 16 214
 FT PROPEP 215 >232
 FT LIPID 214 214
 FT DISULFID 163 198
 FT CARBOHYD 165 165
 FT CARBOHYD 181 181
 FT DOMAIN 44 84
 FT REPEAT 44 51
 FT REPEAT 52 59
 FT REPEAT 60 67
 FT REPEAT 68 75
 FT NON_TER 232 232
 SQ SEQUENCE 232 AA; 25596 MW; 0E2D75F04C05CC4A CRC64;

Query Match 43.1%; Score 50; DB 1; Length 232;
 Best Local Similarity 57.1%; Pred. No. 8.9; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 4 HPGWGKGHPGQGW 17
 | | | | | | | | | |
 DB 53 HGGWGQPHGGGW 66

RESULT 12
 PRIO_CERAT STANDARD; PRT; 238 AA.
 AC Q95145; Q95200;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C) (Fragment).
 GN PRNP.
 OS Cercopithecus aethiops, and
 OS Macaca sylvanus (Barbary ape).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Cercopithecus.
 OX NCBI_TaxID=36222, 9546;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA der Kuyil A.C., Dekker J.T., Goudsmit J.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
 CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
 CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
 CC "RODS".
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
 CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
 CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
 CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
 CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.

CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; U75384; AAB50623.1; -
 DR EMBL; U75382; AAB50629.1; -
 DR HSSP; P04925; IAG2.
 DR InterPro; IPR000817; Prion.
 DR Pfam; PF00377; prion; 1.
 DR SMART; SM00157; PRP; 1.
 DR PROSITE; PS00291; PRION_1; 1.
 DR PROSITE; PS00706; PRION_2; 1.
 KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
 FT NON_TER 1
 FT SIGNAL <1 15
 FT CHAIN 16 215
 FT PROPEP 216 238
 FT LIPID 215 215
 FT DISULFID 164 199
 FT CARBOHYD 166 166
 FT CARBOHYD 182 182
 FT DOMAIN 44 76
 FT REPEAT 44 52
 FT REPEAT 53 60
 FT REPEAT 61 68
 FT REPEAT 69 76
 FT REPEAT 76 76
 SQ SEQUENCE 238 AA; 26123 MW; 5F59A3EBC3E3531B CRC64;

Query Match 43.1%; Score 50; DB 1; Length 238;
 Best Local Similarity 57.1%; Pred. No. 9.2;
 Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 4 HPGWGKGHPGQGW 17
 | | | | | | | | | |
 DB 54 HGGWGQPHGGGW 67

RESULT 13
 PRIO_THEGE STANDARD; PRT; 238 AA.
 AC Q95270;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C) (Fragment).
 GN PRNP OR PRP
 OS Theropithecus gelada (Gelada baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Theropithecus.
 OX NCBI_TaxID=9565;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA der Kuyil A.C., Dekker J.T., Goudsmit J.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
 CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
 CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
 CC "RODS".
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
 CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
 CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
 CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
 CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.

CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U75383; AAB50630.1; -
 CC HSSP: P04925; IAG2.
 CC InterPro: IPR000817; Prion.
 CC Pfam: PF00377; prion; 1.
 CC SMART: SM00157; PRP; 1.
 CC PROSITE: PS00291; PRION_1; 1.
 CC PROSITE: PS00706; PRION_2; 1.
 CC Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
 KW NON_TER 1 1
 FT SIGNAL <1 15
 FT CHAIN 16 >238
 FT DISULFID 164 199
 FT CARBOHYD 166 166
 FT CARBOHYD 182 182
 FT DOMAIN 44 83
 FT REPEAT 44 52
 FT REPEAT 53 60
 FT REPEAT 61 68
 FT REPEAT 69 76
 FT NON_TER 238 238
 SQ SEQUENCE 238 AA; 26104 MW; 5F59BFF602243EDB CRC64;

Query Match 43.1%; Score 50; DB 1; Length 238;
 Best Local Similarity 57.1%; Pred. No. 9.2;
 Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 4 HPGWGQGGHPQGWG 17
 | | | | |
 Db 54 HGGWGQPGHGGWG 67

RESULT 14
 PRIO_AOTTR
 ID PRIO_AOTTR STANDARD; PRT; 239 AA.
 AC P40245;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C) (Fragment).
 GN PRNP.
 OS Aotus trivirgatus (Night monkey) (Douroucoul).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Aotinae; Aotus.
 OX NCBI_TaxID=9505;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95139066; PubMed=7837269;
 RA Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;
 RT "Prion protein gene variation among primates.";
 RL J. Mol. Biol. 245:362-374(1995).
 CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
 CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
 CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
 CC "RODS".
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
 CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
 CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
 CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
 CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
 CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U08293; AAC50082.1; -
 CC HSSP: P04925; IAG2.
 CC InterPro: IPR000817; Prion.
 CC Pfam: PF00377; prion; 1.
 CC SMART: SM00157; PRP; 1.
 CC PROSITE: PS00291; PRION_1; 1.
 CC PROSITE: PS00706; PRION_2; 1.
 CC Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
 KW NON_TER 1 1
 FT SIGNAL <1 15
 FT CHAIN 16 >239
 FT DISULFID 171 206
 FT CARBOHYD 173 173
 FT CARBOHYD 189 189
 FT DOMAIN 44 83
 FT REPEAT 44 51
 FT REPEAT 52 59
 FT REPEAT 60 67
 FT REPEAT 68 75
 FT REPEAT 76 83
 FT NON_TER 239 239
 SQ SEQUENCE 239 AA; 26246 MW; 2EFB77E354B7024A CRC64;

Query Match 43.1%; Score 50; DB 1; Length 239;
 Best Local Similarity 57.1%; Pred. No. 9.2;
 Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 4 HPGWGQGGHPQGWG 17
 | | | | |
 Db 53 HGGWGQPGHGGWG 66

RESULT 15
 PRIO_CALMO
 ID PRIO_CALMO STANDARD; PRT; 241 AA.
 AC P40248;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C) (Fragment).
 GN PRNP.
 OS Callipecus moloch (Dusky titl).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Callipecinae;
 OC Callipecus.
 OX NCBI_TaxID=9523;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95139066; PubMed=7837269;
 RA Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;
 RT "Prion protein gene variation among primates.";
 RL J. Mol. Biol. 245:362-374(1995).
 CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
 CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
 CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
 CC "RODS".
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
 CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
 CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
 CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
 CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.

CC -!- SIMILARITY: BELONGS TO THE PRION FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U08312; AAC50100.1; -
DR HSP; P04925; IAG2.
DR InterPro; IPR000817; Prion.
DR Pfam; PF00377; prion; 1.
DR SMART; SM00157; PRP; 1.
DR PROSITE; PS00291; PRION_1; 1.
DR PROSITE; PS00706; PRION_2; 1.
KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
FT NON_TER 1 1
FT SIGNAL <1 15
FT CHAIN 16 >241
FT DISULFID 172 207
FT CARBOHYD 174 174
FT CARBOHYD 190 190
FT DOMAIN 44 84
FT REPEAT 44 52
FT REPEAT 53 60
FT REPEAT 61 68
FT REPEAT 69 76
FT REPEAT 77 84
FT NON_TER 241 241
SQ SEQUENCE 241 AA; 26373 MW; C6D2013EE7CAEC93 CRC64;

BY SIMILARITY.
MAJOR PRION PROTEIN.
BY SIMILARITY.
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
5 X 8 AA TANDEM REPEATS OF P-H-G-G-W-G-Q.

Query Match 43.18; Score 50; DB 1; Length 241;
Best Local Similarity 57.18; Pred. No. 9.3;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 4 HPQGWGGHPQGWG 17
| | | | |
Db 54 HGGGWGPHGGGWG 67

Search completed: August 6, 2002, 10:40:52
Job time: 249 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 10:40:28 ; Search time 58.49 Seconds
(without alignments)
50.281 Million cell updates/sec

Title: US-09-543-188A-34

Perfect score: 116

Sequence: 1 DGGHPQGGGHPQGWG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL19.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phage.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed. and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	66	56.9	693	16	Q989F7		Q989F7 rhizobium l
2	62	53.4	136	5	Q18444		Q18444 caenorhabdi
3	61.5	53.0	129	5	Q19977		Q19977 caenorhabdi
4	61	52.6	343	10	Q9FUS3		Q9fjs3 arabidopsis
5	60	51.7	156	10	Q39754		Q39754 fagus sylv
6	60	51.7	184	5	Q09943		Q09943 caenorhabdi
7	57	49.1	350	2	Q9XV88		Q9xv88 streptomyc
8	56.5	48.7	473	11	Q91WG7		Q91wg7 mus musculu
9	56.5	48.7	1001	4	Q9H2K7		Q9h2k7 homo sapien
10	56.5	48.7	1001	4	Q96L75		Q96l75 homo sapien
11	56.5	48.7	1001	11	Q88664		Q88664 rattus norv
12	56.5	48.7	1005	4	Q9P216		Q9p216 homo sapien
13	56	48.3	423	10	Q9FM47		Q9fm47 arabidopsis
14	55	47.4	524	4	Q9NXG3		Q9nxg3 homo sapien
15	55	47.4	590	11	Q91WT8		Q91wt8 mus musculu
16	55	47.4	592	6	Q9XSR3		Q9xsr3 canis famil

17	54	46.6	181	5	Q9VWM4	Q9vwm4 drosophila
18	54	46.6	212	10	Q949F9	Q949f9 oryza sativ
19	54	46.6	441	10	Q93YR3	Q93yr3 arabidopsis
20	53	45.7	124	10	Q943P5	Q943p5 oryza sativ
21	53	45.7	137	5	Q9VEH7	Q9veh7 drosophila
22	53	45.7	144	2	Q53360	Q53360 rhizobium m
23	53	45.7	228	11	Q9DLN2	Q9dlm2 mus musculu
24	53	45.7	236	5	Q9TZQ8	Q9tzq8 strongyloce
25	53	45.7	236	5	Q9TZQ7	Q9tzq7 strongyloce
26	53	45.7	247	5	Q9BHA4	Q9bha4 strongyloce
27	53	45.7	250	5	Q9BKN4	Q9bkn4 strongyloce
28	53	45.7	252	5	Q9BKM8	Q9bkm8 strongyloce
29	53	45.7	255	5	Q9BKM9	Q9bkm9 strongyloce
30	53	45.7	256	5	Q9BKN1	Q9bkn1 strongyloce
31	53	45.7	262	5	Q9BKN0	Q9bkn0 strongyloce
32	53	45.7	270	5	Q9TZQ6	Q9tzq6 strongyloce
33	53	45.7	270	5	Q9E536	Q9e536 strongyloce
34	53	45.7	270	5	Q9BKN2	Q9bkn2 strongyloce
35	53	45.7	277	5	Q9BKN3	Q9bkn3 strongyloce
36	53	45.7	302	10	Q9SL09	Q9sl09 arabidopsis
37	53	45.7	325	5	Q9VYM9	Q9vym9 drosophila
38	53	45.7	477	3	Q9HF50	Q9hfe0 orpinomyces
39	52.5	45.3	751	5	Q01497	Q01497 caenorhabdi
40	52	44.8	86	10	Q9LW29	Q9llw29 pinus taeda
41	52	44.8	166	16	Q97M64	Q97m64 clostridium
42	52	44.8	477	3	Q13333	Q13333 orpinomyces
43	52	44.8	557	10	Q9M3C5	Q9m3c5 arabidopsis
44	52	44.8	1292	3	Q96WL0	Q96wl0 ustilago ma
45	51.5	44.4	92	10	Q24350	Q24350 silene lati

ALIGNMENTS

RESULT 1

Q989F7	Q989F7	PRELIMINARY;	PRT;	693 AA.
ID	Q989F7	PRELIMINARY;	PRT;	693 AA.
AC	Q989F7;			
DC	01-OCT-2001 (TREMBLrel. 18, Created)			
DT	01-OCT-2001 (TREMBLrel. 18, Last sequence update)			
DT	01-OCT-2001 (TREMBLrel. 18, Last annotation update)			
DE	ABC TRANSPORTER, PERIPLASMIC ALPHA-GALACTOSIDE BINDING			
DE	PROTEIN.			
GN	MLR6444.			
OS	Rhizobium loti (Mesorhizobium loti).			
OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;			
OC	Phyllobacteriaceae; Mesorhizobium.			
OX	NCBI_TaxID=381;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MAFF303099;			
RX	MEDLINE=21082930; PubMed=11214968;			
RA	Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,			
RA	Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,			
RA	Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,			
RA	Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,			
RA	Takeuchi C., Yamada M., Tabata S.;			
RT	"Complete genome structure of the nitrogen-fixing symbiotic bacterium			
RT	Mesorhizobium loti";			
RL	DNA Res. 7:331-338(2000).			
DR	EMBL; AP003009; BAB52740.1; -			
DR	InterPro; IPR000914; SBP_bac_5.			
DR	Pfam; PF00496; SBP_bac_5; 3.			
KW	Complete proteome.			
SQ	SEQUENCE 693 AA; 77158 MW; 0C8DB281CD4B8ACB CRC64;			

Query Match 56.9%; Score 66; DB 16; Length 693;

Best Local Similarity 61.1%; Pred. No. 1.5;

Matches 11; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

Qy 2 GGHFQGW--GGGHFQGWG 17

||||:|||||

Db 112 GGRPEGWNYGAGOTOGWG 129

RESULT 2

Q18444
ID Q18444 PRELIMINARY; PRT; 136 AA.
AC Q18444;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HYPOTHETICAL 12.3 KDA PROTEIN.
GN C34D4.11.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Du Z., Le T.T.;
RT "The sequence of C. elegans cosmid C34D4.";
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U58755; AAB0096.1; -;
KW Hypothetical protein.
SQ SEQUENCE 136 AA; 12294 MW; F72778C4EA5648B5 CRC64;

Query Match 53.4%; Score 62; DB 5; Length 136;
Best Local Similarity 62.5%; Pred. No. 0.89;
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 GGHPOGWWGGHPQGW 17
Db 58 GGGSGWGGGGGGWG 73

RESULT 3

Q19977
ID Q19977 PRELIMINARY; PRT; 129 AA.
AC Q19977;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE F32G8.3 PROTEIN.
GN F32G8.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurray A.A.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
investigating biology.";

RL Science 282:2012-2018(1998).
DR EMBL; Z72509; CAA96647.1; -;
SQ SEQUENCE 129 AA; 13889 MW; D49198D88E143F03 CRC64;

Query Match 53.0%; Score 61.5; DB 5; Length 129;
Best Local Similarity 73.3%; Pred. No. 0.98;
Matches 11; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 2 GGHPOGWWGGHPQGW 16
Db 33 GGGPGWGGG-PGW 46

RESULT 4

Q9FJ53
ID Q9FJ53 PRELIMINARY; PRT; 343 AA.
AC Q9FJ53;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE GENOMIC DNA, CHROMOSOME 5, P1 CLONE:MJE4.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RX MEDLINE=98403884; PubMed=9734815;
RA Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
Sequence features of the regions of 1,367,185 bp covered by 19
physically assigned P1 and TAC clones.";
RL DNA Res. 5:203-216(1998).
DR EMBL; AB013393; BAB09308.1; -;
DR InterPro; IPR002952; Eggshell.
DR InterPro; IPR000817; Prion.
DR PRINTS; PR01228; EGGSHELL.
DR PRINTS; PR00341; PRION.
SQ SEQUENCE 343 AA; 31629 MW; 67A282819554120A CRC64;

Query Match 52.6%; Score 61; DB 10; Length 343;
Best Local Similarity 64.7%; Pred. No. 3;
Matches 11; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

Qy 2 GGHPOGW--GGGHPQGW 16
Db 87 GGGGGWGGGGGGGGW 103

RESULT 5

Q39754
ID Q39754 PRELIMINARY; PRT; 156 AA.
AC Q39754;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE GRPFI.
OS Fagus sylvatica (Beechnut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Fagaceae; Fagus.
OX NCBI_TaxID=28930;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SEED;
RX MEDLINE=98145483; PubMed=9484489;
RA Nicolas C., Nicolas G., Rodriguez D.;
RT "Transcripts of a gene, encoding a small GTP-binding protein from

RT Fagus sylvatica, are induced by ABA and accumulated in the embryonic
RL axis of dormant seeds.";
PL Plant Mol. Biol. 36:487-491(1998).
[2]

RP SEQUENCE FROM N.A.

RC TISSUE-SEED;

RX MEDLINE=98182934; PubMed=9522463;

RA Nicolas C., Rodriguez D., Poulsen F., Eriksen E.N., Nicolas G.;

RT "The expression of an abscisic acid-responsive glycine-rich protein

coincides with the level of seed dormancy in Fagus sylvatica.;

PL Plant Cell Physiol. 38:1303-1310(1997).

RL EMBL; X98539; CAA67152.1; -

SQ SEQUENCE 156 AA; 15164 MW; 439CC09409F5510E CRC64;

Query Match 51.7%; Score 60; DB 10; Length 156;

Best Local Similarity 62.5%; Pred. No. 1.8;

Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 GGHPQGGGGHGGG 17

III I:|||| I:

Db 54 GGCGGGGGGGGGG 69

RESULT 6

Q09943

ID Q09943 PRELIMINARY; PRT; 184 AA.

AC Q09943;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE F12A10.1 PROTEIN.

GN F12A10.1.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RX MEDLINE=94150718; PubMed=7906398;

RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,

RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,

RA Craxton M., Dear S., Du Z., Durlin R., Favell A., Fulton L.,

RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,

RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,

RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,

RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,

RA Smauld N., Smith A., Sonhammer E., Staden R., Sulston J.,

RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,

RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;

RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.

elegans.";

RL Nature 368:32-38(1994).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA Geisel C.;

RT "The sequence of C. elegans cosmid F12A10.";

RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA Waterston R.;

RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.

DR EMBL; U28731; AAA68294.1; -

SQ SEQUENCE 184 AA; 20062 MW; EF19A9AC696B175A CRC64;

Query Match

Best Local Similarity 51.7%; Score 60; DB 5; Length 184;

Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGHPQGGGGHGGG 16

III I:|||| I:

Db 125 GGYPGGYGGYGGY 139

RESULT 7

Q9X8V8

ID Q9X8V8 PRELIMINARY; PRT; 350 AA.

AC Q9X8V8;

DT 01-NOV-1999 (TrEMBLrel. 12, Created)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE PUTATIVE INTEGRAL MEMBRANE PROTEIN.

GN SCH35.48C.

OS Streptomyces coelicolor.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

OX NCBI_TaxID=1902;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-A3(2);

RA Oliver K., Harris D.;

RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-A3(2);

RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;

RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN-A3(2);

RX MEDLINE=97000351; PubMed=8843436;

RA Redenbach M., Kleser H.M., Denapante D., Eichner A., Cullum J.,

RA Kinashi H., Hopwood D.A.;

RT "A set of ordered cosmids and a detailed genetic and physical map for

the 8 Mb Streptomyces coelicolor A3(2) chromosome.";

RL Mol. Microbiol. 21:77-96(1996).

DR EMBL; AL078610; CAB44379.1; -

SQ SEQUENCE 350 AA; 36081 MW; CCAA14F48CDEB7E5 CRC64;

Query Match

Best Local Similarity 49.1%; Score 57; DB 2; Length 350;

Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 DGGHPQGGGGHQP 14

III I:|||| I:

Db 116 DGGYGGGGGGG 129

RESULT 8

Q91VG7

ID Q91VG7 PRELIMINARY; PRT; 473 AA.

AC Q91VG7;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE HYPOTHETICAL 55.8 KDA PROTEIN.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-BREAST TUMOR;

RA Strausberg R.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC016522; AAH16522.1; -

KW Hypothetical protein.

SQ SEQUENCE 473 AA; 55766 MW; 06B410BF61E705B7 CRC64;

Query Match

Best Local Similarity 48.7%; Score 56.5; DB 11; Length 473;

Best Local Similarity 57.9%; Pred. No. 15;
Matches 11; Conservative 0; Mismatches 5; Indels 3; Gaps 1;

Qy 2 GGHPOGWG---GGHPQGWG 17
Db 399 GGTPQAWGHPMQGGPQPGW 417

RESULT 9

Q9H2K7 PRELIMINARY; PRT; 1001 AA.
AC Q9H2K7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE STE20-LIKE KINASE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Yustein J.T., Robinson D., Kung H.-J.;
RT "Characterization of a Subfamily of Human STE20-like Kinases that
RT Selectively Activate p38 Through MKK3 and are Regulated via a PP2A-
RT dependent Mechanism."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF263312; AAG38502.1; -
DR HSP; P24941; 1838.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Kinase; Transferase.
FT NON_TER 1001
SQ SEQUENCE 1001 AA; 115957 MW; 7F0E534D3E7E159E CRC64;

Query Match 48.7%; Score 56.5; DB 4; Length 1001;
Best Local Similarity 57.9%; Pred. No. 31;
Matches 11; Conservative 0; Mismatches 5; Indels 3; Gaps 1;

Qy 2 GGHPOGWG---GGHPQGWG 17
Db 927 GGTPQAWGHPMQGGPQPGW 945

RESULT 10

Q96L75 PRELIMINARY; PRT; 1001 AA.
AC Q96L75;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE SERINE/THREONINE KINASE TAO1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Jenkins S.G., D'Andrea R.J., Gamble J.R., Vadas M.A.;
RT "Characterization of human TAO1."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY049015; AAL12217.1; -
KW Kinase.
SQ SEQUENCE 1001 AA; 116070 MW; 82941DEAEADC7651 CRC64;

Query Match 48.7%; Score 56.5; DB 4; Length 1001;
Best Local Similarity 57.9%; Pred. No. 31;
Matches 11; Conservative 0; Mismatches 5; Indels 3; Gaps 1;

Qy 2 GGHPOGWG---GGHPQGWG 17
Db 927 GGTPQAWGHPMQGGPQPGW 945

RESULT 11

O88664 PRELIMINARY; PRT; 1001 AA.
AC O88664;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE SERINE/THREONINE PROTEIN KINASE TAO1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99003202; PubMed=9786855;
RA Hutchison M., Berman K.S., Cobb M.H.;
RT "Isolation of TAO1, a protein kinase that activates MEKs in stress-
RT activated protein kinase cascades."
RL J. Biol. Chem. 273:28625-28632(1998).
CC 1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF084205; AAG71014.1; -
DR HSP; P24941; 1838.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1001 AA; 115952 MW; 85511B62DBD62FCC CRC64;

Query Match 48.7%; Score 56.5; DB 11; Length 1001;
Best Local Similarity 57.9%; Pred. No. 31;
Matches 11; Conservative 0; Mismatches 5; Indels 3; Gaps 1;

Qy 2 GGHPOGWG---GGHPQGWG 17
Db 927 GGTPQAWGHPMQGGPQPGW 945

RESULT 12

Q9P216 PRELIMINARY; PRT; 1005 AA.
AC Q9P216;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE KIAA1361 PROTEIN (FRAGMENT).
GN KIAA1361.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20181126; PubMed=10718198;
RA Nagase T., Kikuno R., Ishikawa K., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res. 7:65-73(2000).

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 10:38:48 ; Search time 29.87 Seconds
(without alignments)
22.311 Million cell updates/sec

Title: INVERSE-SEQ23
Perfect score: 36
Sequence: 1 kiwffi 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :			
A_Geneseq_032802.*			
1:	/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT.*		
2:	/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT.*		
3:	/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT.*		
4:	/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT.*		
5:	/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT.*		
6:	/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT.*		
7:	/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT.*		
8:	/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT.*		
9:	/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT.*		
10:	/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT.*		
11:	/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT.*		
12:	/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT.*		
13:	/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT.*		
14:	/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT.*		
15:	/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT.*		
16:	/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT.*		
17:	/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT.*		
18:	/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT.*		
19:	/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT.*		
20:	/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT.*		
21:	/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT.*		
22:	/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	34	94.4	100	AA012383	Human polypeptide
2	33	91.7	345	AB94567	Human protein, sequ
3	33	91.7	972	ABG06365	Novel human diagno
4	33	91.7	1707	ABG22165	Novel human diagno
5	32	88.9	75	AA010644	Human polypeptide
6	32	88.9	102	AA007991	Human polypeptide
7	32	88.9	302	ABG69215	Drosophila melanog
8	32	88.9	627	AA029814	Arabidopsis thalia
9	32	88.9	633	ABG63331	Drosophila melanog
10	32	88.9	842	AA444359	P. chrysogenum sut
11	31	86.1	62	ABG26608	Novel human diagno

12	31	86.1	96	22	AA078726	Human bromine grou
13	31	86.1	110	22	AA081680	S. epidermidis ope
14	31	86.1	119	20	AAV19700	SEQ ID NO 418 from
15	31	86.1	123	20	AAV19814	B. burgdorferi ant
16	31	86.1	136	22	ABG25006	Novel human diagno
17	31	86.1	327	22	ABG64391	Drosophila melanog
18	31	86.1	460	21	AA054132	Human pancreatic c
19	30	83.3	29	22	AA075552	Human secreted pro
20	30	83.3	40	21	AA057017	Arabidopsis thalia
21	30	83.3	69	22	AA006854	Human polypeptide
22	30	83.3	87	20	AA078889	Human secreted pro
23	30	83.3	101	22	AA086857	Human immune/haema
24	30	83.3	106	22	AA064476	Propionibacterium
25	30	83.3	110	19	AAV11054	H. pylori ORF 02ae
26	30	83.3	112	21	AA01617	Human secreted pro
27	30	83.3	136	22	AA075606	Human secreted pro
28	30	83.3	173	21	AA095731	Cosmid CHRM5 enco
29	30	83.3	262	22	ABG65208	Drosophila melanog
30	30	83.3	351	19	AAV11055	H. pylori ORF 07ee
31	30	83.3	497	22	AA062441	Human Acyl CoA rec
32	30	83.3	514	21	AA043167	Human ORFX ORF2931
33	30	83.3	515	22	AA010993	Human lipid metabo
34	30	83.3	515	22	AA092848	Human protein sequ
35	30	83.3	515	22	AA093374	Human protein sequ
36	30	83.3	518	22	AA093470	Human protein sequ
37	30	83.3	518	22	ABG60654	Drosophila melanog
38	30	83.3	519	22	AA039988	Human polypeptide
39	30	83.3	753	22	AA050093	GB1 protein, Caen
40	30	83.3	840	22	AA086159	D. melanogaster GA
41	30	83.3	851	21	AA029627	Cat flea HMT VG Cl
42	30	83.3	1713	22	AB070362	Drosophila melanog
43	30	83.3	1713	22	AA038981	Drosophila G-prote
44	30	83.3	2429	22	ABG62451	Drosophila melanog
45	29	80.6	29	22	ABG06320	Novel human diagno

ALIGNMENTS

RESULT 1
AA012383
ID AA012383 standard; Protein: 100 AA.
XX AC AA012383;
XX DT 06-NOV-2001 (first entry)
XX DE Human polypeptide; SEQ ID NO 26275.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.

XX Homo sapiens.
XX WO200164835-A2.
XX PD 07-SEP-2001.
XX PF 26-FEB-2001; 2001WO-US04927.
XX PR 28-FEB-2000; 2000US-0515126.
XX PR 18-MAY-2000; 2000US-0577409.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX DR WPI; 2001-514838/56.
XX DR N-PSDB; AA192314.
XX PT Isolated nucleic acids and polypeptides, useful for preventing

PT diagnosing and treating e.g. leukaemia, inflammation and immune disorders -

XX

PS Claim 20; SEQ ID NO 26275; 1399pp + Sequence Listing; English.

XX

CC The invention relates to human polynucleotides (AAI79941-AAI93841) and the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation.

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 100 AA;

Query Match 94.4%; Score 34; DB 22; Length 100;

Best Local Similarity 83.3%; Pred. No. 43;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 kiwffi 6

Db 56 kiwffi 61

RESULT 2

AAI94567

ID AAI94567 standard; Protein; 345 AA.

XX

AC AAI94567;

XX

DT 26-JUN-2001 (first entry)

XX

DE Human protein sequence SEQ ID NO:15347.

XX

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX

OS Homo sapiens.

XX

PN EP1074617-A2.

XX

PD 07-FEB-2001.

XX

PF 28-JUL-2000; 2000EP-0116126.

XX

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX

PA (HELI-) HELIX RES INST.

XX

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX

DR WPI; 2001-318749/34.

XX

PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -

XX

PS Claim 8; SEQ ID 15347; 2537pp + CD ROM; English.

XX

CC The present invention describes primer sets for synthesising 5602

CC full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAI03166 to AAI13628 and AAI13633 to AAI18742 represent human cDNA sequences; AAI92446 to AAI95893 represent human amino acid sequences; and AAI13629 to AAI13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

XX

SQ Sequence 345 AA;

Query Match 91.7%; Score 33; DB 22; Length 345;

Best Local Similarity 66.7%; Pred. No. 2.2e+02;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 kiwffi 6

Db 200 kiwffi 205

RESULT 3

ABG06365

ID ABG06365 standard; Protein; 972 AA.

XX

AC ABG06365;

XX

DT 13-FEB-2002 (first entry)

XX

DE Novel human diagnostic protein #6356.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.

XX

OS Homo sapiens.

XX

PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX

PR 30-MAR-2001; 2001WO-US08631.

XX

PR 31-MAR-2000; 2000US-0540217.

XX

PR 23-AUG-2000; 2000US-0649167.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Drmanac RT, Liu C, Tang YT;

XX

DR WPI; 2001-639362/73.

XX

DR N-PSDB; AAS70552.

XX

PT New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity -

XX

PS Claim 20; SEQ ID No 36724; 103pp; English.

XX

CC The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 972 AA;

Query Match 91.7%; Score 33; DB 22; Length 972;
 Best Local Similarity 66.7%; Pred. No. 6.2e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 kiwffi 6
 Db 537 kwifl 542

RESULT 4
 ABG22165
 ID ABG22165 standard; Protein; 1707 AA.
 XX AC ABG22165;
 XX DT 18-FEB-2002 (first entry)
 XX DE Novel human diagnostic protein #22156.
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX OS Homo sapiens.
 XX PN WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US08631.
 XX PR 31-MAR-2000; 2000US-0540217.
 XX PR 23-AUG-2000; 2000US-0649167.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX DR WPI; 2001-639362/73.
 XX DR N-PSDB; AAS86352.
 XX PT New isolated polynucleotide and encoded polypeptides, useful in
 XX PT diagnostics, forensics, gene mapping, identification of mutations
 XX PT responsible for genetic disorders or other traits and to assess
 XX PT biodiversity.
 XX PS Claim 20; SEQ ID NO 52524; 103pp; English.
 XX CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1707 AA;

Query Match 91.7%; Score 33; DB 22; Length 1707;
 Best Local Similarity 66.7%; Pred. No. 1.1e+03;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 kiwffi 6
 Db 537 kwifl 542

RESULT 5
 AAO10644
 ID AAO10644 standard; Protein; 75 AA.
 XX AC AAO10644;
 XX DT 06-NOV-2001 (first entry)
 XX DE Human polypeptide; SEQ ID NO 24536.
 XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 XX KW nervous system disorders; arthritis; inflammation.
 XX OS Homo sapiens.
 XX PN WO200164835-A2.
 XX PD 07-SEP-2001.
 XX PF 26-FEB-2001; 2001WO-US04927.
 XX PR 28-FEB-2000; 2000US-0515126.
 XX PR 18-MAY-2000; 2000US-0577409.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Tang YT, Liu C, Drmanac RT;
 XX DR WPI; 2001-514838/56.
 XX DR N-PSDB; AAI90575.
 XX PT Isolated nucleic acids and polypeptides, useful for preventing
 XX PT diagnosing and treating e.g. leukaemia, inflammation and immune
 XX PT disorders.
 XX PS Claim 20; SEQ ID NO 24536; 1399pp + Sequence Listing; English.
 XX CC The invention relates to human polynucleotides (AAI9941-AAI93841) and
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to

CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 75 AA;

Query Match 88.9%; Score 32; DB 22; Length 75;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 kiwff 5
 |||||
 Db 52 kiwff 56

RESULT 6
 AAC07991
 ID AAC07991 standard; Protein; 102 AA.
 XX
 AC AAC07991;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 21883.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200164835-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 26-FEB-2001; 2001WO-US04927.
 XX
 PR 28-FEB-2000; 2000US-0515126.
 PR 18-MAY-2000; 2000US-0577409.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI; 2001-514838/56.
 DR N-PSDB; AA187922.
 XX
 PT Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosis and treating e.g. leukaemia, inflammation and immune
 PT disorders -
 XX
 PS Claim 20; SEQ ID NO 21883; 1399pp + Sequence Listing; English.
 XX
 CC The invention relates to human polynucleotides (AA179941-AA193841) and
 CC the encoded proteins (AA00010-AA013910) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 102 AA;

Query Match 88.9%; Score 32; DB 22; Length 102;
 Best Local Similarity 100.0%; Pred. No. 98;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 kiwff 5
 |||||
 Db 49 kiwff 53

RESULT 7
 ABB69215
 ID ABB69215 standard; Protein; 302 AA.
 XX
 AC ABB69215;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 34437.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL13318.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 34437; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABB01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 302 AA;

Query Match 88.9%; Score 32; DB 22; Length 302;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 kiwiff 5
Db 132 kiwiff 136

RESULT 8
AAG29814
ID AAG29814 standard; Protein; 627 AA.
XX AC AAG29814;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 35536.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EPI033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159325.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 88.9%; Score 32; DB 21; Length 627;

Best Local Similarity 66.7%; Pred. No. 5e+02; 0; Indels 0; Gaps 0;

Matches 4; Conservative 2; Mismatches 0;

Qy 1 kiwffi 6

Db 20 kwffm 25

RESULT 9

ABB63531
ID ABB63531 standard; Protein; 633 AA.
XX
AC ABB63531;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 17385.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL07634.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
genes from Drosophila and for elucidating cell signalling and cell-cell
interactions -
XX
PS Disclosure; SEQ ID NO 17385; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
capable of detecting 1000 or more genes from Drosophila. The invention is
useful in developmental biology and in elucidating cell signalling and
cell-cell interactions in higher eukaryotes for the development of
insecticides, therapeutics and pharmaceutical drugs. The invention
discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
sequences (ABL01840-ABL16175) and the encoded proteins
sequences (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 633 AA;

Query Match 88.9%; Score 32; DB 22; Length 633;

Best Local Similarity 100.0%; Pred. No. 6.1e+02; 0; Indels 0; Gaps 0;

Matches 5; Conservative 0; Mismatches 0;

Qy 1 kiwff 5

Db 387 kiwff 391

RESULT 10

AA44359

ID AA44359 standard; Protein; 842 AA.

XX

AC AA44359;

XX

DT 14-MAR-2000 (first entry)

XX

DE P. chrysogenum subB protein sequence.

XX

KW subB gene; sulphate permease; sulphur transporter; expression cassette;

KW S-containing beta lactam precursor; beta lactam producing organism;

KW beta lactam compound; penicillin.

```
XX OS Penicillium chrysogenum.
XX PH Key Location/Qualifiers
XX FT Domain 126..147
XX FT /note= "Sulphate permease motif"
XX PN WO9967390-A1.
XX XX XX XX XX
XX PD 29-DEC-1999.
XX XX 22-JUN-1999; 99WO-EP04341.
XX PF 23-JUN-1998; 98EP-0202114.
XX PR (STAM ) DSM NV.
XX PA
XX PI Bovenberg RAL, Van Den Berg MA, Konings WN, Driessen AJM, Turner G;
XX PI Van De Kamp M;
XX XX WPI; 2000-106295/09.
XX DR N-PSDB; AA229467.
XX XX
XX PT New polynucleotide used to improve production of beta-lactam compound,
XX PT and in process of producing the sulphur-containing compound -
XX PS Claim 8; Page 33-36; 47pp; English.
XX CC
XX CC The present sequence is the sulphate transporter encoded by subB gene
XX CC comprising a sulphate permease for sulphate uptake across the cellular
XX CC membrane. This controls the flux of S-containing beta lactam precursors.
XX CC An expression cassette containing the polynucleotide operably linked to a
XX CC regulatory sequence can be used to transform a beta lactam producing
XX CC organism. Transformed cells show an increased production of the
XX CC beta lactam compound. This is used in the control of the regulation and
XX CC expression level of sulphate transporters, in order to influence the
XX CC production level of penicillin and other beta lactams.
XX SQ Sequence 842 AA;

Query Match 88.9%; Score 32; DB 21; Length 842;
Best Local Similarity 86.7%; Pred. No. 8.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 kiwffi 6
Db 291 kmwffv 296
|:|:|:

RESULT 11
ABG26608
ID ABG26608 standard; Protein; 62 AA.
XX AC
XX AC ABG26608;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #26599.
XX DE
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX XX
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX XX
XX PF 30-MAR-2001; 2001WO-US08631.
XX XX
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
```

```
XX (HYSE-) HYSEQ INC.
XX PA
XX PI Drmanac RT, Liu C, Tang YT;
XX XX
XX WPI; 2001-639362/73.
XX DR N-PSDB; AAS90795.
XX XX
XX XX New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -
XX PS Claim 20; SEQ ID No 56967; 103pp; English.
XX CC
XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG0010-ABG30377 represent novel human
XX CC diagnostic amino acid sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 62 AA;

Query Match 86.1%; Score 31; DB 22; Length 62;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 iwffi 6
Db 2 iwffi 6
|:|:|:

RESULT 12
AAG78726
ID AAG78726 standard; Protein; 96 AA.
XX AC
XX AC AAG78726;
XX DT 11-DEC-2001 (first entry)
XX DE Human bromine group domain 11.
XX DE
XX KW Human; bromine group domain 11; cancer; haemopathy; HIV infection;
XX KW immunological disease; inflammation; gene therapy.
XX OS Homo sapiens.
XX XX
XX PN WO200173009-A1.
XX PD 04-OCT-2001.
XX XX
XX PF 26-MAR-2001; 2001WO-CN00424.
XX XX
XX PR 27-MAR-2000; 2000CN-0115152.
XX XX
XX PA (SHAN-) SHANGHAI BIOWINDOW GENE DEV INC.
```

XX PI Mao Y, Xie Y;
 XX WPI; 2001-602862/68.
 DR N-PSDB; AAI66558.
 XX New human bromo group domain 11 for diagnosing and treating malignant
 PT tumor, hemopathy, human immunodeficiency virus infection, immunological
 PT diseases and various inflammations
 XX
 PS Claim 1; Page 30; 35pp; Chinese.
 XX The present invention provides the protein and coding sequences of human
 CC bromine group domain 11. The sequences can be used in the treatment of
 CC cancer, haemopathy, HIV infection, immunological diseases and
 CC inflammation. The present sequence is the protein of the invention.
 XX
 SQ Sequence 96 AA;
 Query Match 86.1%; Score 31; DB 22; Length 96;
 Best Local Similarity 80.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 kiwff 5
 Db 30 kwff 34
 RESULT 13
 AAG81680
 ID AAG81680 standard; Protein: 110 AA.
 AC AAG81680;
 XX 03-SEP-2001 (first entry)
 DT S. epidermidis open reading frame protein sequence SEQ ID NO:454.
 DE Staphylococcus epidermidis SRI strain; infection; diagnosis;
 XX Staphylococcus epidermidis.
 KW vaccination; endocarditis.
 XX
 OS Staphylococcus epidermidis.
 XX
 XX WO200134809-A2.
 PN 17-MAY-2001.
 PD
 XX 09-NOV-2000; 2000WO-US30782.
 EF
 XX 09-NOV-1999; 99US-0164258.
 XX (GLAX) GLAXO GROUP LTD.
 PA
 XX Kimmerly WJ;
 PI
 XX WPI; 2001-316495/33.
 DR N-PSDB; AAH52530.
 XX
 PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 PT useful for vaccinating against infections, e.g. endocarditis -
 XX
 PS Claim 18; Page 160; 2188pp; English.
 XX
 CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
 CC (I) and (II) can have antibacterial activity and therefore can be used
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their

CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH5090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH55091 to
 CC AAH5098 represent oligonucleotide sequences and primers which are used
 CC in the exemplification of the present invention.
 CC N.B. The present invention specifically claims all the polynucleotide
 CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.
 XX
 SQ Sequence 110 AA;
 Query Match 86.1%; Score 31; DB 22; Length 110;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 iwff 6
 Db 100 iwff 104
 RESULT 14
 AAY19700
 ID AAY19700 standard; Protein: 119 AA.
 XX
 AC AAY19700;
 XX
 DT 14-JUL-1999 (first entry)
 XX
 DE SEQ ID NO 418 from WO9922243.
 XX
 KW Human secreted protein; cancer; tumour; neurodegenerative disorder;
 KW developmental abnormality; fetal deficiency; blood disease; leukemia;
 KW immune system disease; autoimmune disease; hepatic disorder;
 KW renal disease; lymphoma; inflammation; allergy; ischemic shock;
 KW Alzheimer's; cognitive disorder; schizophrenia; prostate disease;
 KW obesity; osteoclast; osteoporosis; arthritis; malignancy; testes disease;
 KW lung disease; thymus disease; digestive disorder; endocrine disorder;
 KW infection; AIDS.
 XX
 OS Homo sapiens.
 XX
 XX WO9922243-A1.
 PN
 XX 06-MAY-1999.
 PD
 XX 23-OCT-1998; 98WO-US22376.
 PF
 XX 24-OCT-1997; 97US-0063387.
 PR 24-OCT-1997; 97US-0062784.
 PR 24-OCT-1997; 97US-0063088.
 PR 24-OCT-1997; 97US-0063089.
 PR 24-OCT-1997; 97US-0063090.
 PR 24-OCT-1997; 97US-0063091.
 PR 24-OCT-1997; 97US-0063092.
 PR 24-OCT-1997; 97US-0063097.
 PR 24-OCT-1997; 97US-0063098.
 PR 24-OCT-1997; 97US-0063099.
 PR 24-OCT-1997; 97US-0063100.
 PR 24-OCT-1997; 97US-0063101.
 PR 24-OCT-1997; 97US-0063109.
 PR 24-OCT-1997; 97US-0063110.
 PR 24-OCT-1997; 97US-0063111.
 PR 24-OCT-1997; 97US-0063148.
 PR 24-OCT-1997; 97US-0063386.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Brewer LA, Carter KC, Duan DR, Ebner R, Endress GA;
 PI Feng P, Florence C, Florence KA, Greene JM, Janat F;

PI KayW H, Lafleur DW, Moore PA, Ni J, Olsen HS, Rosen CA;
 PI Ruben SM, Shi Y, Soppet DR, Wei Y, Young P;
 XX WPI; 1999-303069/25.
 DR
 XX
 PT New isolated human genes and the secreted polypeptides they encode
 XX
 PS Disclosure: Page 501-502; 546pp; English.
 XX
 CC The specification describes human secreted proteins. The polynucleotides
 CC and their corresponding secreted polypeptides are useful for preventing,
 CC treating or ameliorating medical conditions, e.g. by protein or gene
 CC therapy. Pathological conditions can also be diagnosed by determining
 CC the amount of the polypeptides in a sample or by determining the presence
 CC of mutations in the polynucleotides. Specific uses are described for each
 CC of the polynucleotides, based on which tissues they are most highly
 CC expressed in, and include developing products for the diagnosis or
 CC treatment of cancer, tumors, neurodegenerative disorders, developmental
 CC abnormalities and fetal deficiencies, blood disorders, leukemias,
 CC diseases of the immune system, autoimmune diseases, hepatic and renal
 CC disease, lymphomas, inflammation, allergies, ischemic shock, Alzheimer's
 CC and cognitive disorders, schizophrenia, prostate diseases, obesity,
 CC disorders involving osteoclasts such as osteoporosis, arthritis or
 CC malignancies, diseases of testes, lung or thymus, digestive/endocrine
 CC disorders, infections and AIDS. The polypeptides are also useful for
 CC identifying their binding partners.
 XX
 SQ Sequence 119 AA;

Query Match 86.1%; Score 31; DB 20; Length 119;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 iwffl 6
 Db 21 iwffl 25

RESULT 15

AAV19814
 ID AAY19814 standard; Protein: 123 AA.

AC AAY19814;

DT 19-JUL-1999 (first entry)

DE B. burgdorferi antigenic protein, f952.aa.

KW Antigenic protein; vaccine; Lyme disease; infection; detection.

OS Borrelia burgdorferi.

PN WO9859071-A1.

XX 30-DEC-1998.

XX 18-JUN-1998; 98WO-US12718.

XX 03-SEP-1997; 97US-0057483.

PR 20-JUN-1997; 97US-0050359.

PR 22-JUL-1997; 97US-0053344.

PR 22-JUL-1997; 97US-0053377.

XX (HUMA-) HUMAN GENOME SCI. INC.

PA (MEDI-) MEDIMUNE INC.

XX Choi GH, Erwin AL, Hanson MS, Lathigra R;

XX WPI; 1999-189980/16.

DR N-PSDB; AAX61511.

XX New isolated Borrelia burgdorferi nucleic acids - used to develop

PT products for the diagnosis, prevention and treatment of diseases
 PT caused by Borrelia, particularly Lyme disease
 XX
 PS Claim 12; Page 75; 275pp; English.

XX This sequence represents a Borrelia burgdorferi (Bb) protein of the
 CC invention, which is suitable for use in a vaccine. The Bb polypeptides
 CC can be used in vaccines for eliciting protective antibodies to members of
 CC the Borrelia genus, particularly for the use against Lyme disease in
 CC humans and animals. They can be used for preventing or attenuating an
 CC infection caused by a member of the Borrelia genus. The products can also
 CC be used for detection of members of the Borrelia genus.

SQ Sequence 123 AA;

Query Match 86.1%; Score 31; DB 20; Length 123;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 iwffl 6

Db 17 iwffl 21

Search completed: August 6, 2002, 10:41:28
 Job time: 160 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 10:38:53 ; Search time 13.02 seconds
(without alignments)
11.256 Million cell updates/sec

Title: INVERSE-SEQ23

Perfect score: 36

Sequence: 1 kiwffi 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PCRUS_COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	80.6	63	4	US-09-188-930-176
2	29	80.6	333	2	US-08-812-871-1
3	29	80.6	333	4	US-09-221-456-2
4	29	80.6	333	4	US-09-558-740-2
5	29	80.6	578	1	US-08-653-740-3
6	29	80.6	578	2	US-09-073-594-3
7	29	80.6	578	3	US-09-275-925-3
8	29	80.6	636	1	US-08-653-740-5
9	29	80.6	636	2	US-09-073-594-5
10	29	80.6	636	3	US-09-275-925-5
11	29	80.6	680	1	US-07-674-287B-2
12	29	80.6	680	2	US-08-436-900A-2
13	29	80.6	682	2	US-08-436-900A-4
14	29	80.6	785	3	US-08-374-077C-3
15	29	80.6	785	4	US-08-895-590-3
16	29	80.6	1720	2	US-08-477-451-12
17	29	80.6	2516	3	US-08-374-077C-2
18	29	80.6	2516	4	US-08-895-590-2
19	29	80.6	3033	1	US-07-925-695-8
20	29	80.6	3033	1	US-07-925-695-9
21	28	77.8	118	4	US-09-627-376-17
22	28	77.8	177	2	US-08-700-013B-11
23	28	77.8	177	2	US-08-700-013B-13
24	28	77.8	186	4	US-09-230-637-36
25	28	77.8	334	2	US-08-700-013B-9
26	28	77.8	471	4	US-08-961-083-158
27	28	77.8	492	1	US-07-794-393-4

Sequence 4, Appli
Sequence 2, Appli
Sequence 10, Appli
Sequence 2, Appli
Sequence 10, Appli
Sequence 19, Appli
Sequence 21, Appli
Sequence 2, Appli
Sequence 27, Appli
Sequence 332, App
Sequence 332, App
Sequence 332, App
Sequence 422, App
Sequence 322, App
Sequence 3, Appli
Sequence 109, App
Sequence 8, Appli
Sequence 92, Appli
Sequence 118, App

ALIGNMENTS

RESULT 1
US-09-188-930-176
; Sequence 176, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Rat
US-09-188-930-176

Query Match 80.6%; Score 29; DB 4; Length 63;
Best Local Similarity 66.7%; Pred. No. 93;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 kiwffi 6
Db 27 KCWFFL 32

RESULT 2
US-08-812-871-1
; Sequence 1, Application US/08812871
; Patent No. 5955303
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl
; APPLICANT: Muzong, Cheng
; TITLE OF INVENTION: NOVEL HUMAN CHEMOKINE RECEPTOR-LIKE
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA

ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/812.871
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0237 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 333 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: MMLR3DT01
CLONE: 568987
US-08-812-871-1

Query Match 80.6%; Score 29; DB 2; Length 333;
Best Local Similarity 80.0%; Pred. NO. 4.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 iwffl 6
Db 146 iwffl 150

RESULT 3
US-09-221-456-2
; Sequence 2, Application US/09221456
; Patent No. 6162899
; GENERAL INFORMATION:
; APPLICANT: SATHE, GANESH
; APPLICANT: HALSEY, WENDY
; APPLICANT: MUIR, ALISON
; APPLICANT: CHAMBERS, JON
; APPLICANT: SZEKERES, PHILIP
; TITLE OF INVENTION: METHODS OF SCREENING FOR AGONISTS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; STREET: P.O. Box 980
; CITY: Valley Forge
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221.456
FILING DATE: 28-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/956,975
FILING DATE: 23-OCT-1997

ATTORNEY/AGENT INFORMATION:
NAME: Prestia, Paul F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70318-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0700
TELEX: 846169
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 333 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-221-456-2

Query Match 80.6%; Score 29; DB 4; Length 333;
Best Local Similarity 80.0%; Pred. NO. 4.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 iwffl 6
Db 146 iwffl 150

RESULT 4
US-09-558-740-2
; Sequence 2, Application US/09558740
; Patent No. 6358695
; GENERAL INFORMATION:
; APPLICANT: SATHE, GANESH
; APPLICANT: HALSEY, WENDY
; APPLICANT: MUIR, ALISON
; APPLICANT: CHAMBERS, JON
; APPLICANT: SZEKERES, PHILIP
; TITLE OF INVENTION: METHODS OF SCREENING FOR AGONISTS AND
; FILE REFERENCE: GH-70318-2
; CURRENT APPLICATION NUMBER: US/09/558,740
; CURRENT FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 08/956,975
; PRIOR FILING DATE: 1997-10-23
; PRIOR APPLICATION NUMBER: 09/221,456
; PRIOR FILING DATE: 1998-12-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 333
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-558-740-2

Query Match 80.6%; Score 29; DB 4; Length 333;
Best Local Similarity 80.0%; Pred. NO. 4.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 iwffl 6
Db 146 iwffl 150

RESULT 5
US-08-653-740-3
; Sequence 3, Application US/08653740
; Patent No. 5792850
; GENERAL INFORMATION:
; APPLICANT: James W. Baumgartner
; APPLICANT: Donald C. Foster
; APPLICANT: Frank J. Grant
; APPLICANT: Cindy A. Sprecher

;; TITLE OF INVENTION: HEMATOPOIETIC CYTOKINE RECEPTOR
;; NUMBER OF SEQUENCES: 42
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: ZymoGenetics, Inc.
;; STREET: 1201 Eastlake Avenue East
;; CITY: Seattle
;; STATE: WA
;; COUNTRY: USA
;; ZIP: 98102
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: IBM PC compatible
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/653,740
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Parker, Gary E
;; REGISTRATION NUMBER: 31,648
;; REFERENCE/DOCKET NUMBER: 95-31
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 206-442-6673
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 578 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-653-740-3

Query Match 80.6%; Score 29; DB 1; Length 578;
Best Local Similarity 50.0%; Pred. No. 7.5e+02;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 kiwffi 6
|:|:|:
Db 267 KWFWV 272

RESULT 6
US-09-073-594-3
;; Sequence 3, Application US/09073594
;; Patent No. 5925735
;; GENERAL INFORMATION:
;; APPLICANT: James W. Baumgartner
;; APPLICANT: Donald C. Foster
;; APPLICANT: Frank J. Grant
;; APPLICANT: Cindy A. Sprecher
;; TITLE OF INVENTION: HEMATOPOIETIC CYTOKINE RECEPTOR
;; NUMBER OF SEQUENCES: 42
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: ZymoGenetics, Inc.
;; STREET: 1201 Eastlake Avenue East
;; CITY: Seattle
;; STATE: WA
;; COUNTRY: USA
;; ZIP: 98102
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: IBM PC compatible
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/073,594
;; FILING DATE:
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Parker, Gary E
;; REGISTRATION NUMBER: 31,648

;; REFERENCE/DOCKET NUMBER: 95-31
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 206-442-6673
;; TELEFAX: 206-442-6678
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 578 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-09-073-594-3

Query Match 80.6%; Score 29; DB 2; Length 578;
Best Local Similarity 50.0%; Pred. No. 7.5e+02;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 kiwffi 6
|:|:|:
Db 267 KWFWV 272

RESULT 7
US-09-275-925-3
;; Sequence 3, Application US/09275925
;; Patent No. 6080406
;; GENERAL INFORMATION:
;; APPLICANT: James W. Baumgartner
;; APPLICANT: Donald C. Foster
;; APPLICANT: Frank J. Grant
;; APPLICANT: Cindy A. Sprecher
;; TITLE OF INVENTION: HEMATOPOIETIC CYTOKINE RECEPTOR
;; NUMBER OF SEQUENCES: 42
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: ZymoGenetics, Inc.
;; STREET: 1201 Eastlake Avenue East
;; CITY: Seattle
;; STATE: WA
;; COUNTRY: USA
;; ZIP: 98102
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: IBM PC compatible
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA: US/09/275,925
;; FILING DATE:
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Parker, Gary E
;; REGISTRATION NUMBER: 31,648
;; REFERENCE/DOCKET NUMBER: 95-31
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 206-442-6673
;; TELEFAX: 206-442-6678
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 578 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-09-275-925-3

Query Match 80.6%; Score 29; DB 3; Length 578;
Best Local Similarity 50.0%; Pred. No. 7.5e+02;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 kiwffi 6
|:|:|:
Db 267 KWFWV 272

RESULT 8
US-08-653-740-5
; Sequence 5, Application US/08653740
; Patent No. 5792850
; GENERAL INFORMATION:
; APPLICANT: James W. Baumgartner
; APPLICANT: Donald C. Foster
; APPLICANT: Frank J. Grant
; APPLICANT: Cindy A. Sprecher
; TITLE OF INVENTION: HEMATOPOIETIC CYTOKINE RECEPTOR
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/653,740
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31,648
; REFERENCE/DOCKET NUMBER: 95-31
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 5:
; LENGTH: 636 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-653-740-5

Query Match 80.6%; Score 29; DB 1; Length 636;
Best Local Similarity 50.0%; Pred. No. 8.2e+02;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 kiwff1 6
I:|:::
Db 267 KWFVW 272

RESULT 9
US-09-073-594-5
; Sequence 5, Application US/09073594
; Patent No. 5925735
; GENERAL INFORMATION:
; APPLICANT: James W. Baumgartner
; APPLICANT: Donald C. Foster
; APPLICANT: Frank J. Grant
; APPLICANT: Cindy A. Sprecher
; TITLE OF INVENTION: HEMATOPOIETIC CYTOKINE RECEPTOR
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,594
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31,648
; REFERENCE/DOCKET NUMBER: 95-31
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 636 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-073-594-5

Query Match 80.6%; Score 29; DB 2; Length 636;
Best Local Similarity 50.0%; Pred. No. 8.2e+02;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 kiwff1 6
I:|:::
Db 267 KWFVW 272

RESULT 10
US-09-275-925-5
; Sequence 5, Application US/09275925
; Patent No. 6080406
; GENERAL INFORMATION:
; APPLICANT: James W. Baumgartner
; APPLICANT: Donald C. Foster
; APPLICANT: Frank J. Grant
; APPLICANT: Cindy A. Sprecher
; TITLE OF INVENTION: HEMATOPOIETIC CYTOKINE RECEPTOR
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/275,925
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31,648
; REFERENCE/DOCKET NUMBER: 95-31
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 636 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-275-925-5

Query Match 80.68; Score 29; DB 3; Length 636;
 Best Local Similarity 50.08; Pred. No. 8.2e+02;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 kiwffi 6
 :|||:
 Db 267 KWEFW 272

RESULT 11
 US-07-674-287B-2
 ; Sequence 2, Application US/07674287B
 ; Patent No. 5414076
 ; GENERAL INFORMATION:
 ; APPLICANT: Bryan Mark O'Hara
 ; TITLE OF INVENTION: Gibbon Ape Leukemia
 ; TITLE OF INVENTION: Virus Receptor
 ; NUMBER OF SEQUENCES: 3
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dr. Karen A. Lowney
 ; ADDRESSEE: American Cyanamid Company
 ; STREET: 1937 West Main Street
 ; STREET: P.O. Box 60
 ; CITY: Stamford
 ; STATE: CT
 ; COUNTRY: USA
 ; ZIP: 06904-0060
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy Disk
 ; COMPUTER: IBM PC AT
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: ASCII converted from IBM DW4
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/674,287B
 ; FILING DATE: 19910325
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lowney, Karen A., Dr.
 ; REGISTRATION NUMBER: 31,274
 ; REFERENCE/DOCKET NUMBER: 31,104-01
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 203 321 2361
 ; TELEFAX: 203 321 2971
 ; TELEX: 710 474 4059
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 680 Amino Acid Residues
 ; TYPE: AMINO ACID SEQUENCE
 ; STRANDEDNESS: Single
 ; TOPOLOGY: Linear
 ; MOLECULE TYPE: Protein
 ; US-07-674-287B-2

Query Match 80.68; Score 29; DB 1; Length 680;
 Best Local Similarity 60.08; Pred. No. 8.7e+02;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 2 iwffi 6
 :|||:
 Db 246 WMEFW 250

RESULT 12
 US-08-436-900A-2
 ; Sequence 2, Application US/08436900A
 ; Patent No. 5874264
 ; GENERAL INFORMATION:

APPLICANT: O'Hara, Bryan M.
 TITLE OF INVENTION: Gibbon Ape Leukemia Virus Receptors
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: American Home Products
 STREET: One Campus Drive
 CITY: Parsippany
 STATE: New Jersey
 COUNTRY: U.S.A.
 ZIP: 07054

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA: US/08/436,900A
 APPLICATION NUMBER: US/08/436,900A
 FILING DATE: 08-MAY-1995
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Barnhard, Elizabeth M.
 REGISTRATION NUMBER: 31,088
 REFERENCE/DOCKET NUMBER: 31,104-03
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-683-2158
 TELEFAX: 201-683-4117
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 680 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-436-900A-2

Query Match 80.68; Score 29; DB 2; Length 680;
 Best Local Similarity 60.08; Pred. No. 8.7e+02;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 2 iwffi 6
 :|||:
 Db 246 WMEFW 250

RESULT 13
 US-08-436-900A-4
 ; Sequence 4, Application US/08436900A
 ; Patent No. 5874264
 ; GENERAL INFORMATION:
 ; APPLICANT: O'Hara, Bryan M.
 ; TITLE OF INVENTION: Gibbon Ape Leukemia Virus Receptors
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: American Home Products
 ; STREET: One Campus Drive
 ; CITY: Parsippany
 ; STATE: New Jersey
 ; COUNTRY: U.S.A.
 ; ZIP: 07054

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/436,900A
 FILING DATE: 08-MAY-1995
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Barnhard, Elizabeth M.
 REGISTRATION NUMBER: 31,088
 REFERENCE/DOCKET NUMBER: 31,104-03

TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-683-2158
TELEFAX: 201-683-4117
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 682 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-436-900A-4

Query Match 80.6%; Score 29; DB 2; Length 682;
Best Local Similarity 60.0%; Pred. No. 8.7e+02;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 kiwffl 6
Db 250 VWFV 254

RESULT 14
US-08-374-077C-3
Sequence 3, Application US/08374077C
Patent No. 6027912
GENERAL INFORMATION:
APPLICANT: Hall, Linda M.
APPLICANT: Ren, Dejlan
APPLICANT: Zheng, Wei
APPLICANT: Dubald, Manuel Marcel Paul
TITLE OF INVENTION: Genes Encoding an Invertebrate Alpha
TITLE OF INVENTION: Calcium Channel Subunit
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, LLP
STREET: 699 Prince Street
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22314-3187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/374,077C
FILING DATE: 19-JAN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm M.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 022650-264
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6620
TELEFAX: 703-836-2021
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 785 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-374-077C-3

Query Match 80.6%; Score 29; DB 3; Length 785;
Best Local Similarity 50.0%; Pred. No. 1e+03;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 kiwffl 6
Db 349 KVMFV 354

RESULT 15
US-08-895-590-3
Sequence 3, Application US/088955590
Patent No. 6207410
GENERAL INFORMATION:
APPLICANT: Hall, Linda M.
APPLICANT: Ren, Dejlan
APPLICANT: Zheng, Wei
APPLICANT: Dubald, Manuel Marcel Paul
TITLE OF INVENTION: Genes Encoding an Insect Calcium Channel
TITLE OF INVENTION: Genes Encoding an Insect Calcium Channel
NUMBER OF SEQUENCES: 101
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, LLP
STREET: 699 Prince Street
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22314-3187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/895,590
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/374,888
FILING DATE: 19-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm M.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 022650-263
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6620
TELEFAX: 703-836-2021
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 785 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-895-590-3

Query Match 80.6%; Score 29; DB 4; Length 785;
Best Local Similarity 50.0%; Pred. No. 1e+03;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 kiwffl 6
Db 349 KVMFV 354

Search completed: August 6, 2002, 10:41:47
Job time: 174 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 10:39:23 ; Search time 14.92 Seconds
(without alignments)
38.642 Million cell updates/sec

Title: INVERSE-SEQ23

Perfect score: 36

Sequence: 1 kiwffi 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	97.2	620	2 F82449	potassium uptake p
2	33	91.7	191	1 W6WLR1	E6 protein - rhesu
3	33	91.7	812	2 AC2349	hypothetical prote
4	32	88.9	556	2 T49501	hypothetical prote
5	32	88.9	565	2 T29813	hypothetical prote
6	31	86.1	125	2 H70101	gIeE protein (gIeE
7	31	86.1	222	2 G72297	conserved hypothet
8	31	86.1	254	2 B90421	hypothetical prote
9	31	86.1	264	2 B72411	conserved hypothet
10	31	86.1	282	2 C91222	hypothetical prote
11	31	86.1	282	2 G86068	hypothetical prote
12	31	86.1	300	2 T16255	hypothetical prote
13	31	86.1	332	2 T33310	hypothetical prote
14	31	86.1	336	1 S75272	cytochrome d ubiq
15	31	86.1	463	2 T19343	hypothetical prote
16	31	86.1	523	2 S50479	28S proteasome reg
17	31	86.1	551	2 S64314	probable membrane
18	31	86.1	552	2 T39121	amino-acid permea
19	31	86.1	680	2 T42923	infected cell prot
20	31	86.1	1039	2 T15885	hypothetical prote
21	31	86.1	2352	2 T43431	alpha-glucan synth
22	30	83.3	90	2 AH0464	probable membrane
23	30	83.3	131	2 AH0348	probable membrane
24	30	83.3	162	2 CH9838	conserved hypothet
25	30	83.3	170	2 A64347	conserved hypothet
26	30	83.3	172	2 S55015	NADH dehydrogenase
27	30	83.3	175	2 B69011	conserved hypothet
28	30	83.3	220	2 S78395	ribosomal protein
29	30	83.3	293	2 B83783	hypothetical prote

30 83.3 313 2 T28728
31 30 83.3 336 2 A86406
32 30 83.3 340 2 B84019
33 30 83.3 351 2 D71955
34 30 83.3 373 2 E84595
35 30 83.3 390 1 ACBPMG
36 30 83.3 403 2 A83344
37 30 83.3 405 2 T19593
38 30 83.3 452 2 G95347
39 30 83.3 456 2 T40416
40 30 83.3 468 2 S10014
41 30 83.3 493 2 G90604
42 30 83.3 536 2 S66716
43 30 83.3 569 2 D69374
44 30 83.3 749 2 T41072
45 30 83.3 809 2 S40460

ALIGNMENTS

RESULT 1

F82449

Potassium uptake protein, Kup system VCA0529 [imported] - Vibrio cholerae (strain N16

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: F82449

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers

l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833

A:Accession: F82449

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-620 <HEI>

A:Cross-references: G3:AE004383; GB:AE003853; NID:g9657927; PIDN:AAF96432.1; GSPDB:GN

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VCA0529

A:Map position: 2

Query Match 97.2%; Score 35; DB 2; Length 620;

Best Local Similarity 83.3%; Pred. No. 57;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0

Qy 1 kiwffi 6

Db 586 KWFFI 591

RESULT 2

W6WLR1

E6 protein - rhesus papillomavirus (type 1)

C:Species: rhesus papillomavirus

C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Feb-1997

C:Accession: A38503

R:Ostrow, R.S.; LaBresh, K.V.; Faras, A.J.

Virology 181, 424-429, 1991

A:Title: Characterization of the complete RHPV 1 genomic sequence and an integration

A:Reference number: A38503; MUID:91135018

A:Accession: A38503

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-191 <OST>

A:Cross-references: EMBL:M37717

C:Superfamily: papillomavirus E6 protein

C:Keywords: DNA binding; early protein; transforming protein; zinc finger

F:60-96/Region: zinc finger CCCC motif

F:133-169/Region: zinc finger CCCC motif

A:Molecule type: DNA
A:Residues: 1-222 <ARN>
A:Cross-references: GB:AE001768; GB:AE000512; NID:g4981619; PIDN:AAD36164.1; PID:g498163
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM1087

Query Match 86.1%; Score 31; DB 2; Length 222;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 iwffi 6
:||||:
Db 177 IWFFI 181

RESULT 8
B90421
hypothetical protein SS02489 [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C:Accession: B90421
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
arratt, R.A.; Ragan, M.A.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: B90421
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-254 <KUR>
A:Cross-references: GB:AE006641; NID:g13815792; PIDN:AAK42625.1; GSPDB:GN00155
C:Genetics:
A:Gene: SS02489

Query Match 86.1%; Score 31; DB 2; Length 254;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 iwffi 6
:||||:
Db 127 IWFFI 131

RESULT 9
B72411
conserved hypothetical protein - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: B72411
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316
A:Accession: B72411
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-264 <ARN>
A:Cross-references: GB:AE001701; GB:AE000512; NID:g4980648; PIDN:AAD35257.1; PID:g498065
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM0164

Query Match 86.1%; Score 31; DB 2; Length 264;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 kiwff 5
:||||:
Db 2 KWFF 6

RESULT 10
C91222
hypothetical protein ECS4747 [imported] - Escherichia coli (strain O157:H7, substrain
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: C91222
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: C91222
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-282 <HA1>
A:Cross-references: GB:BA000007; PIDN:BAR38170.1; PID:g13364223; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: ECS4747

Query Match 86.1%; Score 31; DB 2; Length 282;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 kiwffi 6
:||||:
Db 252 RIWFFL 257

RESULT 11
G86068
hypothetical protein Z5335 [imported] - Escherichia coli (strain O157:H7, substrain E
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: G86068
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A35480; MUID:21074935; PMID:11206551
A:Accession: G86068
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-282 <STQ>
A:Cross-references: GB:AE005174; NID:g12518689; PIDN:AAG59011.1; GSPDB:GN00145; UWGP:
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z5335

Query Match 86.1%; Score 31; DB 2; Length 282;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 kiwffi 6
:||||:
Db 252 RIWFFL 257

RESULT 12
T16255
hypothetical protein F35C8.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Mar-2000
C:Accession: T16255
R:Wu, X.
submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid F35C8.

A:Reference number: Z18486

A:Accession: T16255

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-300 <WUX>

A:Cross-references: EMBL:U40941; NID:gl072184; PID:AAA81710.1; CESP:F35C8

C:Genetics:

A:Gene: CESP:F35C8.5

A:Introns: 32/3: 62/1: 154/3: 184/3: 219/3

C:Superfamily: Saccharomyces cerevisiae ERG25 protein

Query Match 86.1%; Score 31; DB 2; Length 300;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 iwffl 6

|||||

Db 208 IWFFI 212

RESULT 13

T33310

hypothetical protein K02H11.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

A:Accession: T33310

R:Rohlfing, T.; Wohldmann, P.; Antoniou, B.

submitted to the EMBL Data Library, May 1998

A:Description: The sequence of C. elegans cosmid K02H11.

A:Reference number: Z21320

A:Accession: T33310

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-332 <ROH>

A:Cross-references: EMBL:AF068720; PID:AAI17786.1; GSPDB:GN00023; CESP:K02H11.7

A:Experimental source: strain Bristol N2; clone K02H11

C:Genetics:

A:Gene: CESP:K02H11.7

A:Map position: 5

A:Introns: 119/2; 162/2; 314/1

Query Match 86.1%; Score 31; DB 2; Length 332;

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 iwffl 6

|||||

Db 131 IWFFI 135

RESULT 14

S75272

cytochrome d ubiquinol oxidase (EC 1.10.3.-) chain II - Synechocystis sp. (strain PCC 68

N:Alternate names: protein slr1380

C:Species: Synechocystis sp.

A:Variety: PCC 6803

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000

C:Accession: S75272

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;

O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996

A>Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis

s.

A:Reference number: S74322; MUID:97061201

A:Accession: S75272

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-336 <KAN>

A:Cross-references: EMBL:D90904; GB:AB001339; NID:gl652225; PID:BAAL17186.1; PID:gl65226

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Genetics:

A:Gene: cydB

C:Superfamily: cytochrome d ubiquinol oxidase

C:Keywords: electron transfer; heme; oxidoreductase; respiratory chain; transmembrane

Query Match 86.1%; Score 31; DB 1; Length 336;

Best Local Similarity 66.7%; Pred. No. 1.7e+02;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 kiwffi 6

|||||

Db 13 QWFFI 18

RESULT 15

T19343

hypothetical protein Cl7D12.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999

C:Accession: T19343; T26960

R:White, S.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19112

A:Accession: T19343

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-463 <WIL>

A:Cross-references: EMBL:Z81473; PIDN:CA803891.1; GSPDB:GN00019; CESP:Cl7D12.3

A:Experimental source: clone Cl7D12

R:Harris, B.

submitted to the EMBL Data Library, September 1998

A:Reference number: Z20291

A:Accession: T26960

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-463 <WIL>

A:Cross-references: EMBL:AL031636; PIDN:CAA21046.1; GSPDB:GN00019; CESP:Y47H9A.1

A:Experimental source: clone Y47H9A

C:Genetics:

A:Gene: CESP:Cl7D12.3; CESP:Y47H9A.1

A:Map position: 1

A:Introns: 18/2; 61/1; 136/3; 178/3; 202/3; 235/1; 317/3; 362/3; 386/3; 446/3

Query Match 86.1%; Score 31; DB 2; Length 463;

Best Local Similarity 100.0%; Pred. No. 2.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 iwffl 6

|||||

Db 455 IWFFI 459

Search completed: August 6, 2002, 10:42:10

Job time: 167 sec

16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Metaxin 1 homolog.
CG9393.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Beremand B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Maitel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kimons I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith J.C.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -!- FUNCTION: INVOLVED IN TRANSPORT OF PROTEINS INTO THE
CC MITOCHONDRION. ESSENTIAL FOR EMBRYONIC DEVELOPMENT (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Mitochondrial outer membrane (By
CC similarity).
CC -!- SIMILARITY: BELONGS TO THE METAXIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE003693; AAF54402.1; -.
CC Flybase: FBgn0037710; CG9393.
CC InterPro: IPR004046; GST_C.
CC Hypothetical protein; Mitochondrion; Outer membrane; Transmembrane;
KW Transmembrane protein transport.
FT TRANSMEM 281 301 POTENTIAL.
SO SEQUENCE 327 AA: 37252 MW: 83366528AF6580A2 CRC64;

Query Match 86.1%; Score 31; DB 1; Length 327;
Best Local Similarity 80.0%; Pred. No. 91;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Oy 1 kiwiff 5
Db 190 KWIFF 194

RESULT 3
RPN3_YEAST
ID AC P40016;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 26S proteasome regulatory subunit RPN3.
GN RPN3 OR SUN2 OR YER021W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96242146; PubMed=8668124;
RA Kawamura M., Kominami K.-I., Takeuchi J., Toh-E A.;
RT "A multicopy suppressor of nin1-1 of the yeast Saccharomycetes
cerevisiae is a counterpart of the Drosophila melanogaster diphenol
oxidase A2 gene, Dox-A2."
RL Mol. Gen. Genet. 251:146-152(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97170075; PubMed=9017604;
RA Kominami K.-I., Okura N., Kawamura M., Demartino G.N., Slaughter C.A.,
RA Shimbara N., Chung C.H., Fujimuro M., Yokosawa H., Shimizu Y.,
RA Tanahashi N., Tanaka K., Toh-E A.;
RT "Yeast counterparts of subunits S5a and p58 (S3) of the human 26S
proteasome are encoded by two multicopy suppressors of nin1-1."
RL Mol. Biol. Cell 8:171-187(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C / AB972;
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
RA Chung E., Duncan M., Guzman E., Hartzell G., Hunicke-Smith S.,
RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ACTS AS A REGULATORY SUBUNIT OF THE 26 PROTEASOME WHICH
CC IS INVOLVED IN THE ATP-DEPENDENT DEGRADATION OF UBIQUITINATED
CC PROTEINS (BY SIMILARITY).
CC -!- SUBUNIT: THE 26S PROTEASOME IS COMPOSED OF A CORE PROTEASE, KNOWN
CC AS THE 20S PROTEASOME, CAPPED AT ONE OR BOTH ENDS BY THE 19S
CC REGULATORY COMPLEX (RC). THE RC IS COMPOSED OF AT LEAST 18
CC DIFFERENT SUBUNITS IN TWO SUBCOMPLEXES, THE BASE AND THE LTD,
CC WHICH FORM THE PORTIONS PROXIMAL AND DISTAL TO THE 20S PROTEOLYTIC
CC CORE, RESPECTIVELY (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE PROTEASOME SUBUNIT S3 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: D78023; BAAL1208.1; -.
CC EMBL: U18778; AAB64554.1; -.

DR SGD; S0000823; RPN3.
DR InterPro; IPR00717; PCI.
DR Pfam; PF01399; PCI; 1.
DR SMART; SM00088; PINT; 1.

FT	CONFLICT	355	355	S -> G (IN REF. 3).
SQ	SEQUENCE	523	AA; 60422	MW; D0DA1645B8DE958D CRC64;

```
Query Match      86.1%; Score 31; DB 1; Length 523;
Best Local Similarity 66.7%;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 kiwffi 6
|:|:|:
Db 182 KLWFFYI 187

RESULT	4	YG1F_YEAST	STANDARD;	PRT;	551 AA.
ID	YG1F_YEAST	STANDARD;	PRT;	551 AA.	
AC	P53214;				
DT	01-OCT-1996	(Rel. 34, Created)			
DT	01-OCT-1996	(Rel. 34, Last sequence update)			
DT	15-DEC-1998	(Rel. 37, Last annotation update)			
DE	Hypothetical 57.5 kDa protein in VMA7-RPS25A intergenic region.				
GN	YGR023W.				
OS	Saccharomyces cerevisiae (Baker's yeast).				
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;				
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.				
NCBI	TaxID=4932;				

[1] _____, 1997,
RN
SEQUENCE FROM N.A.
RP
RC
MEDLINE=97435481; PubMed=9290212;
RX
Rieder M., Brueckner M., Schaefer M., Mueller-Auer S.;
RT "Sequence analysis of 203 kilobases from *Saccharomyces cerevisiae*.
 chromosome VII.";
RL
yeast 13:1077-1090(1997).
CC
-!- SIMILARITY: TO YEAST MID2.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

[illegible]

Query Match	86.1%	Score 31	DB 1	Length 551
Best Local Similarity	66.7%	Pred. No.	1.4e+02	
Matches 4	Conservative	2	Mismatches	0
			Indels	0
			Gaps	0

Qy 1 kiwffi 6

Db 412 KIWF 417

RESULT	5			
YLK6_CAEEL				
ID	YLK6_CAEEL	STANDARD;	PRT;	1039 AA.
AC	P41954;			
DT	01-NOV-1995	(Rel. 32, Created)		
DT	01-NOV-1995	(Rel. 32, Last sequence update)		
DT	01-NOV-1995	(Rel. 32, Last annotation update)		
DE	Hypochemical 131.8 kDa protein D1044.6 in chromosome III.			
GN	D1044.6			

OS *Caenorhabditis elegans*.
 OS Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderae; Caenorhabditis.
 ON NCBI_TaxID=6239;
 RX [1]

RN
 RP
 RC
 RA
 RL
 CC

[1]
 SEQUENCE FROM N.A.
 STRAIN-BRISTOL N°;
 Pauley A., Waterston R.;
 Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

```

CC      EMBL; U00065; AAA50738.1; *.
DR      WormPep; D1044.6; CE01209.
DR      Hypothetical protein;
KW      SEQUENCE 1039 AA; 121848 MW; 69785B2B360E569F CRC64;
OR      Send an Email to: ncicbncb@nci.nih.gov

```

Query Match 86.1%; Score 31; DB 1; Length 1039;
Best Local Similarity 80.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels

Qy	1	kiwff	5
		:	
Db	5	KVWFF	9

RESULT	.6
MOKC_SCHPO	
ID	MOKC_SCHPO STANDARD; PRT; 2352 AA.
AC	Q9UUL4; O13605;
DT	16-OCT-2001 (Rel. 40, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	Cell wall alpha-1,3-glucan synthase mok12 (EC 2.4.1.183).
GN	MOK12 OR SPEC3ZH8_13C OR PI011.
OS	Schizosaccharomyces pombe (Fission yeast).
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC	Schizosaccharomycetales; Schizosaccharomycetaceae;
OC	Schizosaccharomycetes.
OX	NCBI_TaxID=4896;

RN [1] "Genbank 1989",
RP SEQUENCE FROM N.A.
RC STRAIN-972,
RA Katayama S., Dai H., Arellano M., Perez P., Toda T.,
RT "Fission yeast alpha-glucan synthase Mok1 localizes
RL and play a role essential for cell morphogenesis and
function.",
RN Submitted (OCT-1998) to the EMBL/GenBank/DDBB database

RN
RP
RC
RX
RX
RX
RX
[2]
SEQUENCE FROM N.A.
STRAIN=972;
MEDLINE=20089027; PubMed=10620777;
Machida M., Yamazaki S., Kunihiro S.,
Tanaka T., Kushida N., Jinno K.,

```

RA Haikawa Y., Yamazaki J., Yamamoto S., Sekine M., Oguchi A., Nagai Y.,
RA Sakai M., Aoki K., Ogura K., Kudo Y., Kikuchi H., Zhang M.O.,
RA Yanagida M.;
RT "A 38 kb segment containing the cdc2 gene from the left arm of fission
RT yeast chromosome II: sequence analysis and characterization of the
RT genomic DNA and cDNAs encoded on the segment.";
RL Yeast 16:71-80(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Wood V., Rajandream M.A., Barrell B.G., Seeger K., Harris D.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: UDP-glucose + [alpha-D-glucosyl-(1,3)](N)-
CC UDP + [alpha-D-glucosyl-(1,3)](N+1).
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AB018381; BAA76558.1; -
DR EMBL; AB004534; BAA21388.1; ALT_INIT.
DR EMBL; AL590971; CAC37503.1; -
DR InterPro; IPR000461; Alpha_amyase.
DR InterPro; IPR001296; Glycos_transf_1.
DR Pfam; PF00128; alpha-amyase; 1.
DR Pfam; PF00534; Glycos_transf_1; 1.
DR Cell wall; Transferase; Glycosyltransferase.
KW SEQUENCE 2352 AA; 266561 MW; 78ADF9C2F7140BBA CRC64;
SQ

```

Query Match 86.1%; Score 31; DB 1; Length 2352;
 Best Local Similarity 100.0%; Pred. No. 5.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 iwffi 6
 Db 2145 IWFFI 2149

```

RESULT 7
Y377_METJA
ID Y377_METJA STANDARD; PRT; 170 AA.
AC Q57822;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0377.
GN MJ0377.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurt M.A., Raine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U67490; AAB98366.1; -
DR TIGR; MJ0377; -
DR InterPro; IPR002785; DUF83.
DR Pfam; PF01930; DUF83; 1.
DR PRODom; PD012943; DUF83; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 170 AA; 20420 MW; 10CEBC592550AA4B CRC64;

```

Query Match 83.3%; Score 30; DB 1; Length 170;
 Best Local Similarity 50.0%; Pred. No. 77;
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 kiwffi 6
 Db 25 KLWYFV 30

```

RESULT 8
NU6M_PETMA
ID NU6M_PETMA STANDARD; PRT; 172 AA.
AC Q35544;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 6 (EC 1.6.5.3).
GN MTND6 OR ND6 OR NAD6.
OS Petromyzon marinus (Sea lamprey).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95229067; PubMed=7713438;
RA Lee W.J., Kocher T.D.;
RT "Complete sequence of a sea lamprey (Petromyzon marinus)
RT mitochondrial genome: early establishment of the vertebrate genome
RT organization.";
RL Genetics 139:873-887(1995).
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone -> NAD(+) + ubiquinol.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U11880; AAB08749.1; -
DR InterPro; IPR001457; Oxidored_q3.
DR Pfam; PF00499; oxidored_q3; 1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 172 AA; 18463 MW; C69D941E959B4B52 CRC64;

```

Query Match 83.3%; Score 30; DB 1; Length 172;
 Best Local Similarity 80.0%; Pred. No. 78;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 iwffi 6
 Db 87 IWFFV 91


```
RESULT 9
RR3_EPIVI
ID RR3_EPIVI STANDARD; PRT; 220 AA.
AC P30055;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Chloroplast 30S ribosomal protein S3.
GN RPS3.
OS Epifagus virginiana (Beechdrops).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Lamiales; Orobanchaceae; Epifagus.
OX NCBI_TaxID=4177;
[1]
SEQUENCE FROM N.A.
RR MEDLINE=93066301; PubMed=1332054;
RX Wolfe K.H., Morden C.W., Palmer J.D.;
RA "Function and evolution of a minimal plastid genome from a
RT nonphotosynthetic parasitic plant.";
PL Proc. Natl. Acad. Sci. U.S.A. 89:10648-10652(1992).
[2]
SEQUENCE FROM N.A.
RR MEDLINE=93021155; PubMed=1404416;
RX Wolfe K.H., Morden C.W., Ems S.C., Palmer J.D.;
RA "Rapid evolution of the plastid translational apparatus in a
RT nonphotosynthetic plant: loss or accelerated sequence evolution of
RT tRNA and ribosomal protein genes.";
RL J. Mol. Evol. 35:304-317(1992).
[3]
-1- SIMILARITY: BELONGS TO THE S3P FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M81884; AAA5864.1;
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00189; Ribosomal_S3.C; 1.
DR SMART; SM00417; Ribosomal_S3.N; 1.
DR Pfam; PF00417; Ribosomal_S3.N; 1.
DR PROSITE; PS00548; RIBOSOMAL_S3; 1.
KW Ribosomal protein; Chloroplast.
SQ SEQUENCE 220 AA; 25939 MW; 5BF70AC74AB7DF94 CRC64;
Query Match 83.3%; Score 30; DB 1; Length 220;
Best Local Similarity 83.3%; Pred. No. 96;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 kiwiffi 6
DB 212 KIWI 217
RESULT 10
Y258_HELPJ
ID Y258_HELPJ STANDARD; PRT; 350 AA.
AC Q92M88;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein JHP0242.
GN JHP0242.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
[1]
SEQUENCE FROM N.A.
RR MEDLINE=99120557; PubMed=9923682;
RX Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE YAEI/HI0918/HP0258/SLR1821 FAMILY.
CC -1- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001462; AAD05832.1; ALT_INIT.
DR InterPro; IPR001478; PDZ.
DR Pfam; PF00595; PDZ; 1.
DR SMART; SM00228; PDZ; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 43 63 POTENTIAL.
FT TRANSMEM 94 114 POTENTIAL.
FT TRANSMEM 249 269 POTENTIAL.
FT TRANSMEM 277 297 POTENTIAL.
FT TRANSMEM 326 346 POTENTIAL.
SQ SEQUENCE 350 AA; 38864 MW; 99DCA574282D6AF9 CRC64;
Query Match 83.3%; Score 30; DB 1; Length 350;
Best Local Similarity 80.0%; Pred. No. 1,4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 kiwiff 5
DB 41 KIWIFF 45
RESULT 11
VASS_BPGE
ID VASS_BPGE STANDARD; PRT; 390 AA.
AC P07394;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Assembly protein (Maturation protein) (A protein).
GN A.
OS Bacteriophage GA.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Leviviridae;
OC Levivirus.
OX NCBI_TaxID=12018;
RN [1]
SEQUENCE FROM N.A.
RR MEDLINE=86223910; PubMed=3711059;
RX Inokuchi Y., Takahashi R., Hirose T., Inayama S., Jacobson A.B.,
RA Hirashima A.;
RT "The complete nucleotide sequence of the group II RNA coliphage GA.";
RL J. Biochem. 99:1169-1180(1986).
CC -1- FUNCTION: THE MATURATION PROTEIN IS REQUIRED FOR THE TYPICAL
CC ATTACHMENT OF THE PHAGE TO THE SIDE OF THE BACTERIAL PILI.
CC IT ACCOMPANY THE VIRAL DNA INTO THE CELL.
CC -1- PTM: THE FORMYL METHIONINE MAY BE REMOVED AFTER TRANSLATION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; D10027; BAA00917.1; -
 DR EMBL; X03869; CAA27496.1; -
 DR PIR; JS0009; ACBPMG.
 KW Phage recognition; Formylation.
 FT MOD_RES 1 FORMYLATION
 SQ SEQUENCE 390 AA; 44385 MW; B3B86A05209A1735 CRC64;

Query Match 83.3%; Score 30; DB 1; Length 390;
 Best Local Similarity 66.7%; Pred. No. 1.6e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 kiwffi 6
 DB 262 RIWYFI 267

RESULT 12

GLCP_SYNY3 STANDARD; PRT; 468 AA.
 ID GLCP_SYNY3
 AC P15729;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Glucose transport protein.
 GN GTR OR GLCP OR SLL0771.
 OS Synecocystis sp. (strain PCC 6803).
 CC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
 OX NCBI_TaxID=11148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90014182; PubMed=2507869;
 RA Zhang C.C., Durand M.C., Jeanjean R., Joset F.;
 RT "Molecular and genetical analysis of the fructose-glucose transport
 system in the cyanobacterium Synecocystis PCC6803.";
 RL Mol. Microbiol. 3:1221-1229(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91346660; PubMed=2129397;
 RA Schmetterer G.R.;
 RT "Sequence conservation among the glucose transporter from the
 cyanobacterium Synecocystis sp. PCC 6803 and mammalian glucose
 transporters.";
 RL Plant Mol. Biol. 14:697-706(1990).
 RN [3]

RP SEQUENCE FROM N.A.
 RX MEDLINE=96127529; PubMed=8590279;
 RA Kaneo T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
 RA Sugitani M., Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 Synecocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
 region from map positions 64% to 92% of the genome.";
 RL DNA Res. 2:153-166(1995).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X15988; CAA34119.1; -
 DR EMBL; X16472; CAA34492.1; -
 DR EMBL; D64000; BAA10117.1; -
 DR PIR; S06973; S06973.

PIR; S10014; S10014.
 DR InterPro; IPR003663; Sugar_transporter.
 DR InterPro; IPR003662; sub_transporter.
 DR Pfam; PF00083; sugar_tr_1.
 DR PRINTS; PR00171; SUGRTRANSPORT.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transmembrane; Transport; Sugar transport; Complete proteome.
 FT DOMAIN 1 17 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 18 38 1 (POTENTIAL).
 FT DOMAIN 39 58 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 59 78 2 (POTENTIAL).
 FT DOMAIN 79 84 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 85 105 3 (POTENTIAL).
 FT DOMAIN 106 114 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 115 135 4 (POTENTIAL).
 FT DOMAIN 136 149 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 150 170 5 (POTENTIAL).
 FT DOMAIN 171 186 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 187 207 6 (POTENTIAL).
 FT DOMAIN 208 265 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 266 286 7 (POTENTIAL).
 FT DOMAIN 287 307 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 308 328 8 (POTENTIAL).
 FT DOMAIN 329 331 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 332 352 9 (POTENTIAL).
 FT DOMAIN 353 366 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 367 387 10 (POTENTIAL).
 FT DOMAIN 388 412 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 413 433 11 (POTENTIAL).
 FT DOMAIN 434 436 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 437 457 12 (POTENTIAL).
 FT DOMAIN 458 468 CYTOPLASMIC (POTENTIAL).
 FT CONFLICT 319 319 L -> I (IN REF. 1).
 SQ SEQUENCE 468 AA; 49747 MW; D7EC545C4FB38D22 CRC64;

Query Match 83.3%; Score 30; DB 1; Length 468;
 Best Local Similarity 80.0%; Pred. No. 1.9e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 iwffi 6
 DB 453 IWFFV 457

RESULT 13

SYEM_YEAST STANDARD; PRT; 536 AA.
 ID SYEM_YEAST
 AC P48525; Q08203;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Glutamyl-tRNA synthetase, mitochondrial (EC 6.1.1.17) (Glutamate--tRNA
 DE ligase) (GLURS).
 GN MSEI OR YOLO33W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D273-10B;
 RA Tzagoloff A.A., Shtanko A.;
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Habbig B., Hattenhorst U., Hollenhorst C.P., Ramezani Rad M.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + tRNA(Glu) = AMP +
 CC diphosphate + L-glutamyl-tRNA(Glu).
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-tRNA SYNTHETASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L39015; AAA61403.1; -
DR EMBL: 274775; CAA99033.1; -
DR HSP; P27000; IGLN.
DR SGD; S0005393; NSEI.
DR InterPro: IPR000924; trna-synt_1c.
DR InterPro: IPR001412; trna-synt_1.
DR Pfam: PF00749; trna-synt_1c; 1.
DR PRINTS: PR00987; TRNASYNTHGLU.
DR PROSITE: PS00178; AA-trna_LIGASE_I; FALSE_NEG.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Mitochondrion.
FT SIFE 53 61 "HIGH" REGION.
FT CONFLICT 464 464 F -> L (IN REF. 1).
SQ SEQUENCE 536 AA; 61603 MW; 5CF36FBAD0E8C58C CRC64;

Query Match 83.3%; Score 30; DB 1; Length 536;
Best Local Similarity 80.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 kiwff 5
|:|:|
Db 362 KIWF 366

RESULT 14

ID DDC_CAEEL STANDARD; PRT; 830 AA.
AC P34751; Q23619;
DT 01-FEB-1994 (Rel. 28, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Probable aromatic L-amino-acid decarboxylase (EC 4.1.1.28) (AADC)
DE (DOPA decarboxylase) (DDC).
GN AAD-1 OR ZK829.2
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Harris B.;
RU Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 206-830 FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=93173104; PubMed=8382340;
RA Marra M.A., Prasad S.S., Baillie D.L.;
RT "Molecular analysis of two genes between let-653 and let-56 in the
unc-22(IV) region of Caenorhabditis elegans."
RL Mol. Gen. Genet. 236:289-298(1993).
CC -1- CATALYTIC ACTIVITY: L-TRYPTOPHAN AND DIHYDROXY-L-PHENYLALANINE (DOPA)).
ON 5-HYDROXY-L-TRYPTOPHAN AND L-TRYPTOPHAN = TRYPTAMINE + CO(2) (ALSO ACTS
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: CATECHOLAMINE BIOSYNTHESIS.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
TYRDC).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z73899; CAA98072.1; -
DR EMBL: Z11576; CAA77663.1; -
DR PIR: S30909; S30909.
DR PIR: S19796; S19796.
DR WormPep: ZK829.2; CE066650.
DR InterPro: IPR002129; Pyridoxal_dec.
DR Pfam: PF00282; Pyridoxal_dec; 1.
DR PRINTS: PR00800; YHCRBOXLASE.
DR PROSITE: PS00392; DDC_GAD_HDC_YDC; FALSE_NEG.
KW Lyase; Decarboxylase; Catecholamine biosynthesis; Pyridoxal phosphate.
FT BINDING 573 573 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT CONFLICT 340 347 HPNHSFY -> SSKFSFIL (IN REF. 2).
SQ SEQUENCE 830 AA; 93221 MW; E884D5261567DC89 CRC64;

Query Match 83.3%; Score 30; DB 1; Length 830;
Best Local Similarity 83.3%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 kiwff 6
|:|:|:|
Db 618 KIWFMI 623

RESULT 15

ID OAR_MYXXX STANDARD; PRT; 1061 AA.
AC P38370;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE OAR protein precursor.
GN OAR.
OS Myxococcus xanthus.
OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
OC Myxococcales; Cystobacterineae; Myxococcaceae; Myxococcus.
OX NCBI_TaxID=34;
[1]
RN SEQUENCE FROM N.A., AND SEQUENCE OF 776-796.
RP STRAIN=DZF1;
RX MEDLINE=93328680; PubMed=8335633;
RA Martinez-Canamero M., Munoz-Dorado J., Farex-Vidal E., Inouye M.,
RA Inouye S.;
RT "Oar, a 115-kilodalton membrane protein required for development of
Myxococcus xanthus."
RL J. Bacteriol. 175:4756-4763(1993).
CC -1- FUNCTION: REQUIRED FOR CELLULAR ADHESION DURING FRUITING BODY
FORMATION.
CC -1- SUBCELLULAR LOCATION: OUTER-MEMBRANE ASSOCIATED.
CC -1- SIMILARITY: LOCAL, TO TONB-DEPENDENT RECEPTOR PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S64103; AAB27614.1; -
DR PIR: A40609; A40609.
KW Outer membrane; Signal.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 1061 OAR PROTEIN.
SQ SEQUENCE 1061 AA; 114455 MW; EA8C077296352EFO CRC64;

Query Match 83.3%; Score 30; DB 1; Length 1061;
Best Local Similarity 80.0%; Pred. No. 3.8e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 kiwff 5
|:|:|
Db 300 KLWFF 304

Search completed: August 6, 2002, 10:42:58
Job time: 124 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 10:40:33 ; Search time 25.01 Seconds
(without alignments)
41.502 Million cell updates/sec

Title: INVERSE-SEQ23
Perfect score: 36
Sequence: 1 kiwffi 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protist:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	36	100.0	321	8 Q9G4C0	Q9g4c0 thraustochy
2	35	97.2	620	16 Q9KM59	Q9km59 vibrio chol
3	34	94.4	894	5 Q9NEZ0	Q9nez0 caenorhabdi
4	33	91.7	345	4 Q9H8R0	Q9h8r0 homo sapien
5	33	91.7	1487	4 Q9ULM3	Q9ulm3 homo sapien
6	32	88.9	274	12 Q9DWC2	Q9dwc2 rat cytoameg
7	32	88.9	302	5 Q9VE58	Q9ve58 drosophila
8	32	88.9	565	5 Q18651	Q18651 caenorhabdi
9	32	88.9	633	5 Q9V7Q0	Q9v7q0 drosophila
10	32	88.9	731	5 Q20121	Q20121 caenorhabdi
11	32	88.9	842	3 Q9URR4	Q9urr4 penicillium
12	31	86.1	93	8 Q9G3Z9	Q9g3z9 lithobius f
13	31	86.1	93	8 Q94RE6	Q94re6 lithobius f
14	31	86.1	125	16 O51048	O51048 borrelia bu
15	31	86.1	192	16 Q98NG4	Q98ng4 rhizobium l
16	31	86.1	205	17 Q97BH7	Q97bh7 thermoplasm

17	31	86.1	214	17 Q9HJ12	Q9hj12 thermoplasm
18	31	86.1	222	16 Q9XOH3	Q9xoh3 thermotoga
19	31	86.1	254	17 Q97VW5	Q97vw5 sulfolobus
20	31	86.1	284	16 Q9WY11	Q9wy11 thermotoga
21	31	86.1	300	5 Q20027	Q20027 caenorhabdi
22	31	86.1	312	8 Q9MT25	Q9mi25 drosophila
23	31	86.1	312	8 Q9MI23	Q9mi23 drosophila
24	31	86.1	312	8 Q9MI21	Q9mi21 drosophila
25	31	86.1	312	8 Q9MI19	Q9mi19 drosophila
26	31	86.1	332	5 O61982	O61982 caenorhabdi
27	31	86.1	336	16 P73160	P73160 synechocyst
28	31	86.1	463	5 O62065	O62065 caenorhabdi
29	31	86.1	513	8 Q9MXX6	Q9mux6 lophatherum
30	31	86.1	513	8 Q95F28	Q95f28 enneapogon
31	31	86.1	519	8 Q9MV25	Q9mv25 oryza ridle
32	31	86.1	519	8 Q9MV24	Q9mv24 oryza longi
33	31	86.1	521	8 Q9MD18	Q9md18 scenedesmus
34	31	86.1	523	16 Q9CMX3	Q9cmx3 pasteurella
35	31	86.1	552	3 Q9UR23	Q9ur23 schizosacche
36	31	86.1	560	2 Q9F8U1	Q9f8ul streptomyce
37	31	86.1	560	4 Q9BSJ2	Q9bsj2 homo sapien
38	31	86.1	661	2 O68468	O68468 corynebacte
39	31	86.1	680	12 Q9YQ6	Q9yq6 ateline her
40	31	86.1	699	13 Q73735	Q73735 xenopus lae
41	31	86.1	902	4 O43632	O43632 homo sapien
42	31	86.1	905	11 Q921G8	Q921g8 mus musculu
43	30	83.3	50	4 Q9NRC0	Q9nrc0 homo sapien
44	30	83.3	125	10 Q92P72	Q92p72 schinzella
45	30	83.3	162	16 Q99VU7	Q99vu7 staphylococ

ALIGNMENTS

RESULT 1

Q9G4C0 ID Q9G4C0 PRELIMINARY; PRT; 321 AA.
AC Q9G4C0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE NADH DEHYDROGENASE SUBUNIT 1 (EC 1.6.5.3).
GN NAD1.
OS Thraustochytrium aureum.
OG Mitochondrion.
OC Eukaryota: stramenopiles; Labyrinthulida; Thraustochytriidae;
OC Thraustochytrium.
OX NCBI_TaxID=42467;
RN [1]
RP SEQUENCE FROM N.A.
RA Burger G., Lang B.F., Gray M.W.;
RT "Phylogenetic relationships of stramenopile algae, based on complete mitochondrial genome sequences."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF288091; AAG23680.1; -;
DR InterPro; IPR001694; Resp_chain_NADH_DH1.
DR Pfam; PF00146; NADHdh; 1.
DR PROSITE; PS00668; COMPLEX1_ND1_2; 1.
KW Oxidoreductase; Mitochondrion.
SQ SEQUENCE 321 AA; 36063 MW; 886C2C671756818E CRC64;

Query Match 100.0%; Score 36; DB 8; Length 321;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 kiwffi 6

Db 175 KIWFFI 180

RESULT 2

Q9KM59

ID Q9KM59 PRELIMINARY; PRT; 620 AA.
AC Q9KM59;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE POTASSIUM UPTAKE PROTEIN, KUP SYSTEM.
GN VCA0529.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Emolaleva M.D., Vamathevan J., Bass S., Olin H., Dragol I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
Fraser C.M.;
RA "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae";
RL Nature 406:477-483(2000).
DR EMBL; AE004383; AAF96432.1; -;
DR TIGR; VCA0529; -;
DR InterPro; IPR003855; K+ transprtr.
KW Complete proteome.
SQ SEQUENCE 620 AA; 69032 MW; 840973C674E141C9 CRC64;

Query Match 97.2%; Score 35; DB 16; Length 620;
Best Local Similarity 83.3%; Pred. No. 99;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 kiwffi 6

Db 586 KWFFFI 591

RESULT 3
Q9NEZ0 PRELIMINARY; PRT; 894 AA.
ID Q9NEZ0;
AC Q9NEZ0;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Y105EBE.J PROTEIN (Y105E8A.M PROTEIN).
GN Y105EBE.J OR Y105E8A.M.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Sulston J.E.;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology";
RL Science 282:2012-2018(1998).
RN [3]
RP SEQUENCE FROM N.A.
RA Sulston J.E.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL132880; CAB60873.2; -;
DR EMBL; AL132876; CAC48126.1; -;
DR InterPro; IPR002553; Adaptin.N.
DR InterPro; IPR001121; G_adapt_C.
DR Pfam; PF01602; Adaptin_N; 1.

SQ SEQUENCE 894 AA; 99765 MW; 477B050C252E0586 CRC64;

Query Match 94.4%; Score 34; DB 5; Length 894;
Best Local Similarity 83.3%; Pred. No. 2.le+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 kiwffi 6

Db 599 KIWFFL 604

RESULT 4

Q9H8R0 PRELIMINARY; PRT; 345 AA.
ID Q9H8R0;
AC Q9H8R0;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE OVARC1001436 PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=OVARIAN CARCINOMA;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H., Sugawara M.,
RA Wagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Takiguchi S.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEDO human cDNA sequencing project";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK023370; BAB14546.1; -;
SQ SEQUENCE 345 AA; 38757 MW; D8461EC4AFFFE158 CRC64;

Query Match 91.7%; Score 33; DB 4; Length 345;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 kiwffi 6

Db 200 KWFFFL 205

RESULT 5

Q9ULM3 PRELIMINARY; PRT; 1487 AA.
ID Q9ULM3;
AC Q9ULM3;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE KIAA1197 PROTEIN (FRAGMENT).
GN KIAA1197.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=20039619; PubMed=10574462;
RA Nagase T., Ishikawa K., Kikuno R., Hirose M., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro";
RL DNA Res. 6:337-345(1999).
DR EMBL; AB033023; BAA86511.1; -;
DR NON_TER 1

SQ SEQUENCE 1487 AA; 157165 MW; 9DC4FC22CCCF4414 CRC64;

Query Match 91.7%; Score 33; DB 4; Length 1487;
Best Local Similarity 66.7%; Pred. No. 5.3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 kiwiff 6
|:||||
Db 318 KWFFFL 323

RESULT 6
Q9DWC2 ID Q9DWC2 PRELIMINARY; PRT; 274 AA.

AC Q9DWC2: 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
GN R79.

OS Rat cytomegalovirus (strain Maastricht).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Muromegalovirus.

OX NCBI_TaxID=79700;
[1]

RP SEQUENCE FROM N.A.
RC STRAIN=MAASTRICHT;

RX MEDLINE=99370163; PubMed=10438809;

RA Beisser P.S., Grauls G., Bruggeman C.A., Vink C.;

RT "Deletion of the R78 G protein-coupled receptor gene from rat cytomegalovirus results in an attenuated, syncytium-inducing mutant strain.";

RL J. Virol. 73:7218-7230(1999).

RN [2]
RP SEQUENCE FROM N.A.

RC STRAIN=MAASTRICHT;

RX MEDLINE=20366325; PubMed=10906222;

RA Vink C., Beuken E., Bruggeman C.A.;

RT "Complete DNA sequence of the rat cytomegalovirus genome.";

RL J. Virol. 74:7656-7665(2000).

RN [3]
RP SEQUENCE FROM N.A.

RC STRAIN=MAASTRICHT;

RX MEDLINE=20473137; PubMed=11018281;

RA Gruijthuisen Y.K., Beuken E., Bruggeman C.A., Vink C.;

RT "Rat cytomegalovirus R89 is a highly conserved gene which expresses a spliced transcript.";

RL Virus Res. 69:119-130(2000).

DR EMBL; AF232689; AAF99168.1; -.

SQ SEQUENCE 274 AA; 31287 MW; BB6FADF5AC441978 CRC64;

Query Match 88.9%; Score 32; DB 12; Length 274;
Best Local Similarity 100.0%; Pred. No. 1.0e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 kiwiff 5
|:||||
Db 204 KIWF 208

RESULT 7
Q9VE58 ID Q9VE58 PRELIMINARY; PRT; 302 AA.

AC Q9VE58: 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
GN CG14305 PROTEIN.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;
[1]

RN SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;

RX MEDLINE=20196008; PubMed=10731132;

RA Adams M.D., Celniker S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA Ananides P.G., Scherer S.E., Richards S., Ashburner M., Henderson S.N.,
RA George R.A., Lewis S.E., Yandell M.D., Zhang Q., Chen L.X.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benca P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195(2000).

-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

EMBL; AE003722; AAF55571.1; -.

HSSP; Q63450.1406.

Flybase; FBgn0038630; CG14305.

InterPro; IPR000719; Euk_pkinase.

InterPro; IPR002290; Ser_thr_pkinase.

InterPro; IPR001245; Tyr_pkinase.

Pfam; PF00069; pkinase; 1

PRINTS; PR00109; TYRKINASE.

PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.

PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

KW ATP-binding; Serine/threonine-protein kinase; Transferase.

SQ SEQUENCE 302 AA; 34276 MW; 3055BCD3B0A1495A CRC64;

Query Match 88.9%; Score 32; DB 5; Length 302;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 kiwiff 5
|:||||
Db 132 KIWF 136

RESULT 8

Q18651

ID Q18651 PRELIMINARY; PRT; 565 AA.

AC Q18651;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE HYPOTHETICAL 65.9 KDA PROTEIN.
 GN C46A5.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RL investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Johnson D., Stellyes L.;
 RT "The sequence of C. elegans cosmid C46A5.";
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U61948; AA803149.1; -;
 DR InterPro; IPR001873; ASC.
 DR Pfam; PF00858; ASC; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 565 AA; 65894 MW; 8DAC1078563FE7C3 CRC64;

Query Match 88.9%; Score 32; DB 5; Length 565;
 Best Local Similarity 66.7%; Pred. No. 3.2e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 kiwiffi 6
 I:||||
 DB 63 KIWFFL 68

RESULT 9
 Q9V700 PRELIMINARY; PRT; 633 AA.
 AC Q9V700;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE CG7848 PROTEIN (GH17657P).
 GN CG7848.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA Sutton G.G., Wortman J.R., Richards S., Ashburner M., Henderson S.N.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abriel J.F., Agbayani A., An H.-J., Andrews-pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J., Ye J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhao G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Y, CN BW SP;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Munnall C.J.,
 RA Nuncio J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003807; AAF57997.1; -;
 DR EMBL; AY058405; AAL13634.1; -;
 DR FlyBase; FBgn0034127; CG7848.
 DR InterPro; IPR001891; Malic_enzyme.
 DR Pfam; PF00390; malic; 2.
 DR PRINTS; PR00072; MALOXRDASE.
 SQ SEQUENCE 633 AA; 72362 MW; 6EDB66494DD347E2 CRC64;

Query Match 88.9%; Score 32; DB 5; Length 633;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 kiwiff 5
 I:||||
 DB 387 KIWFF 391

RESULT 10
 Q20121 PRELIMINARY; PRT; 731 AA.
 AC Q20121;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE HYPOTHETICAL 81.3 KDA PROTEIN.
 GN F37C12.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;

RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RL investigating biology. The C. elegans Sequencing Consortium.";
 RN Science 282:2012-2018(1998).
 RP [2]
 RC SEQUENCE FROM N.A.
 RD STRAIN-BRISTOL N2;
 RE Fulton L.;
 RF "The sequence of C. elegans cosmid F37C12.";
 RG Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 RH [3]
 RI SEQUENCE FROM N.A.
 RJ STRAIN-BRISTOL N2;
 RK Waterston R.;
 RL "Direct Submission.";
 RM EMBL; U00033; AAC48292.3; --
 RN InterPro; IPR000873; AMP-bind.
 RO Pfam; PF00501; AMP-binding; 1.
 RP PRINTS; PR00154; AMPBINDING.
 RQ PROSITE; PS00455; AMP_BINDING; 1.
 RW Hypothetical protein.
 RY SEQUENCE 731 AA; 81252 MW; 1F317B9D899C5FC0 CRC64;

Query Match 88.9%; Score 32; DB 5; Length 731;
 Best Local Similarity 66.7%; Pred. No. 4.1e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 kiwffi 6
 Db 36 RWFFFI 41

RESULT 11
 Q9URR4 PRELIMINARY; PRT; 842 AA.
 AC Q9URR4;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE SULFATE PERMEASE SUBT.
 GN SUTB.
 OS Penicillium chrysogenum.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.
 OX NCBI_TaxID=5076;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Q176;
 RX MEDLINE=20042342; PubMed=10572125;
 RA Van de Kamp M., Pizzini E., Vos A., Van der Lende T.R.,
 RA Schuurs T.A., Newbert R.W., Turner G., Konings W.N., Driessen A.J.M.;
 RT "Sulfate Transport in Penicillium chrysogenum: Cloning and
 RT Characterization of the sutA and sutB Genes.";
 RL J. Bacteriol. 181:7228-7234(1999).
 DR EMBL; AF163974; AAF14539.1; --
 DR InterPro; IPR002645; STAS.
 DR InterPro; IPR001902; Sulfate_transp.
 DR Pfam; PF01740; STAS; 1.
 DR Pfam; PF00916; Sulfate_transp; 1.
 DR PROSITE; PS01130; SULFATE_TRANS; UNKNOWN.1.
 SQ SEQUENCE 842 AA; 91865 MW; 839A55486E733D15 CRC64;

Query Match 88.9%; Score 32; DB 3; Length 842;
 Best Local Similarity 66.7%; Pred. No. 4.7e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 kiwffi 6
 Db 291 KMWFFV 296

RESULT 12
 Q9G3Z9 PRELIMINARY; PRT; 93 AA.
 AC Q9G3Z9;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE NADH DEHYDROGENASE SUBUNIT 4L.
 GN NAD4L.
 OS Lithobius forficatus.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Myriapoda; Chilopoda;
 OC Pleurostigmophora; Lithobiomorpha; Lithobiidae; Lithobius.
 OX NCBI_TaxID=7552;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20538576; PubMed=11095730;
 RA Lavrov D.V., Brown W.M., Boore J.L.;
 RT "A novel type of RNA editing occurs in the mitochondrial tRNAs of the
 RT centipede Lithobius forficatus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:13738-13742(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Lavrov D.V., Brown W.M., Boore J.L.;
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF309492; AAG39994.1; --
 DR InterPro; IPR00314; Mit_NADHub_oxidredctse_4L.
 DR InterPro; IPR001133; Oxidored_q2.
 DR Pfam; PF00420; oxidored_q2; 1.
 DR ProDom; PD000359; Mit_NADHub_oxidredctse_4L; 1.
 KW Mitochondrion.
 SQ SEQUENCE 93 AA; 10491 MW; 062CD0A404053B46 CRC64;

Query Match 86.1%; Score 31; DB 8; Length 93;
 Best Local Similarity 100.0%; Pred. No. 86;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 iwffi 6
 Db 43 IWFFI 47

RESULT 13
 Q94RE6 PRELIMINARY; PRT; 93 AA.
 AC Q94RE6;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE NADH DEHYDROGENASE SUBUNIT 4L (FRAGMENT).
 GN ND4L.
 OS Lithobius forficatus.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Myriapoda; Chilopoda;
 OC Pleurostigmophora; Lithobiomorpha; Lithobiidae; Lithobius.
 OX NCBI_TaxID=7552;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hwang U.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21441907; PubMed=11557978;
 RA Hwang U., Friedrich M., Choe C., Kim W.;
 RT "Mitochondrial Protein Phylogeny joins myriapods with chelicerates.";
 RL Nature 413:154-157(2001).
 DR EMBL; AJ270997; CAC69945.1; --
 KW Mitochondrion.
 FT NON_TER 93
 SQ SEQUENCE 93 AA; 10491 MW; FC9016DEB8C7AB5E CRC64;

Query Match 86.1%; Score 31; DB 8; Length 93;
 Best Local Similarity 100.0%; Pred. No. 86;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 iwffl 6
 |||||
 Db 43 IWFFI 47

RESULT 14

OS1048 PRELIMINARY; PRT; 125 AA.
 AC OS1048;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE GLPE PROTEIN (GLPE).
 GN BB0016.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID-139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 35210 / B31;
 RX MEDLINE-98065943; PubMed-9403685;
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
 RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
 RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
 RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
 RA Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
 RA Smith H.O., Venter J.C.;
 RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
 burgdorferi";
 RL Nature 390:580-586(1997).
 DR EMBL; AE001115; AAC66391.1; -.
 DR TIGR; BB0016; -.
 DR InterPro; IPR001763; Rhodanese_domain.
 DR SMART; SM00450; RHOD; 1.
 KW Complete proteome.
 SQ SEQUENCE 125 AA; 14557 MW; 42F7A03D3A365C92 CRC64;

Query Match 86.1%; Score 31; DB 16; Length 125;
 Best Local Similarity 100.0%; Pred. No. 11e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 iwffl 6
 |||||
 Db 19 IWFFI 23

RESULT 15

Q98NG4 PRELIMINARY; PRT; 192 AA.
 AC Q98NG4;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
 DE MLR0150 PROTEIN.
 GN MLR0150.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID-381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MAFF303099;
 RX MEDLINE-21082930; PubMed-11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watarabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,

RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti";
 RL DNA Res. 7:331-338(2000).
 DR EMBL; AP002994; BAB47797.1; -.
 DR InterPro; IPR000086; NUDIX_hydrolase.
 DR Pfam; PF00293; NUDIX; 1.
 KW Complete proteome.
 SQ SEQUENCE 192 AA; 21950 MW; 50063BC9B46DFF34 CRC64;

Query Match 86.1%; Score 31; DB 16; Length 192;
 Best Local Similarity 66.7%; Pred. No. 1.7e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 kiwffl 6
 ::|||
 Db 127 RLWFFI 132

Search completed: August 6, 2002, 10:42:42
 Job time: 129 sec